BnF Gallica

Calcul des probabilités / par J.-B. Pomey,...

Source gallica.bnf.fr / Bibliothèque nationale de France

{BnF Gallica

Pomey, Jean-Baptiste (1861-1943). Auteur du texte. Calcul des probabilités / par J.-B. Pomey,.... 1936.

1/ Les contenus accessibles sur le site Gallica sont pour la plupart des reproductions numériques d'oeuvres tombées dans le domaine public provenant des collections de la BnF. Leur réutilisation s'inscrit dans le cadre de la loi n°78-753 du 17 juillet 1978 :
- La réutilisation non commerciale de ces contenus ou dans le cadre d'une publication académique ou scientifique est libre et gratuite dans le respect de la législation en vigueur et notamment du maintien de la mention de source des contenus telle que précisée ci-après : « Source gallica.bnf.fr / Bibliothèque nationale de France » ou « Source gallica.bnf.fr / BnF ».
- La réutilisation commerciale de ces contenus est payante et fait l'objet d'une licence. Est entendue par réutilisation commerciale la revente de contenus sous forme de produits élaborés ou de fourniture de service ou toute autre réutilisation des contenus générant directement des revenus : publication vendue (à l'exception des ouvrages académiques ou scientifiques), une exposition, une production audiovisuelle, un service ou un produit payant, un support à vocation promotionnelle etc.

CLIQUER ICI POUR ACCÉDER AUX TARIFS ET À LA LICENCE

2/ Les contenus de Gallica sont la propriété de la BnF au sens de l'article L.2112-1 du code général de la propriété des personnes publiques.

3/ Quelques contenus sont soumis à un régime de réutilisation particulier. Il s'agit :

- des reproductions de documents protégés par un droit d'auteur appartenant à un tiers. Ces documents ne peuvent être réutilisés, sauf dans le cadre de la copie privée, sans l'autorisation préalable du titulaire des droits.
- des reproductions de documents conservés dans les bibliothèques ou autres institutions partenaires. Ceux-ci sont signalés par la mention Source gallica.BnF.fr / Bibliothèque municipale de ... (ou autre partenaire). L'utilisateur est invité à s'informer auprès de ces bibliothèques de leurs conditions de réutilisation.

4/ Gallica constitue une base de données, dont la BnF est le producteur, protégée au sens des articles L341-1 et suivants du code de la propriété intellectuelle.

5/ Les présentes conditions d'utilisation des contenus de Gallica sont régies par la loi française. En cas de réutilisation prévue dans un autre pays, il appartient à chaque utilisateur de vérifier la conformité de son projet avec le droit de ce pays.

6/ L'utilisateur s'engage à respecter les présentes conditions d'utilisation ainsi que la législation en vigueur, notamment en matière de propriété intellectuelle. En cas de non respect de ces dispositions, il est notamment passible d'une amende prévue par la loi du 17 juillet 1978.

7/ Pour obtenir un document de Gallica en haute définition, contacter
utilisation.commerciale@bnf.fr.

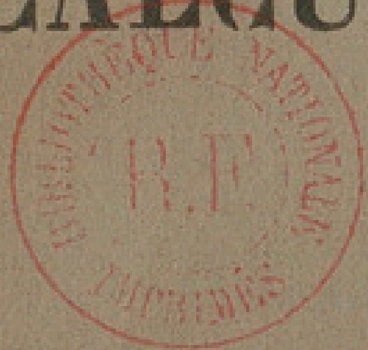
4° V
12865

Source gallica.bnf.fr / Bibliothèque nationale de France

Cart

CONFÉRENCE SUR LE CALCUL DES PROBABILITÉS
A L'ÉCOLE SUPÉRIEURE D'ÉLECTRICITÉ

CALCUL DES PROBABILITÉS



2/235

PAR

J.-B. POMEY
Ingénieur en chef des Postes et Télégraphes



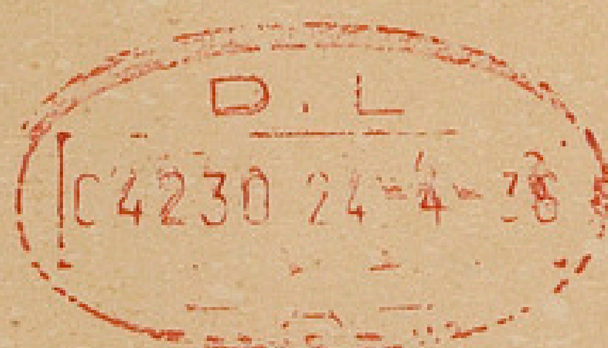
PARIS
GAUTHIER-VILLARS, ÉDITEUR
LIBRAIRE DU BUREAU DES LONGITUDES, DE L'ÉCOLE POLYTECHNIQUE
Quai des Grands-Augustins, 55.

1936

CALCUL DES PROBABILITÉS



4 V
12865



B

Tous droits de traduction et de reproduction réservés pour tous pays.

CALCUL DES PROBABILITÉS

PRINCIPES GÉNÉRAUX.

1. **Notion de probabilité.** — Un problème de probabilité n'est déterminé que si l'on a spécifié les cas que l'on considère comme également probables.

Exemple : on donne une urne contenant des boules toutes pareilles, sauf que les unes sont blanches et les autres noires. On tire une boule de l'urne et l'on admet que toutes les boules, quelle que soit leur couleur, ont même chance d'être tirées.

Il y a a blanches et b noires; si l'on tire une boule de l'urne, quelle est la probabilité pour qu'elle soit blanche?

La réponse est immédiate : il y a $a+b$ façons de tirer une boule, il y a a façons de tirer une blanche. Si donc nous définissons la *probabilité* comme le rapport du nombre des cas favorables au nombre total des cas possibles, quand ceux-ci sont considérés comme également admissibles, la probabilité cherchée sera

$$\frac{a}{a+b}.$$

La probabilité d'amener une noire sera de même

$$\frac{b}{a+b}.$$

Il est certain qu'on amènera soit une blanche, soit une noire, or on a

$$\frac{a}{a+b}+\frac{b}{a+b}=1.$$

La somme des probabilités est égale à 1 et la probabilité 1 est le symbole de la *certitude*.

On amènera soit une blanche, soit une noire. Ces deux événements sont *contradictoires*. Chacun des cas possibles, s'il n'est

pas favorable à l'un est favorable à l'autre. La somme des probabilités de deux événements contradictoires est égale à l'unité.

Ajoutons maintenant à l'urne c boules rouges, ces boules étant identiques aux précédentes, sauf la couleur qu'on ne peut distinguer, à cause de l'urne qui est opaque.

Les probabilités se définiront de la même manière comme le rapport du nombre des cas favorables au nombre des cas possibles, supposés tous également probables.

Les probabilités seront les suivantes :

1° Pour amener une blanche, $\frac{a}{a+b+c}$;

2° Pour amener une noire, $\frac{b}{a+b+c}$;

3° Pour amener une rouge, $\frac{c}{a+b+c}$.

Si l'on amène une blanche, on ne peut amener une noire, les deux événements sont *exclusifs* l'un de l'autre, mais ils ne sont pas contradictoires, car si l'on n'amène pas une blanche, on peut fort bien ne pas amener une noire, et c'est ce qui arrivera, si l'on amène une rouge.

La probabilité d'amener une blanche ou une noire est égale à

$$\frac{a+b}{a+b+c}.$$

C'est, d'autre part, la *somme* des probabilités respectives d'amener une blanche et d'amener une noire. Il n'y a plus que deux catégories de boules, l'une formée de la réunion des blanches et des noires, l'autre formé des rouges.

Dans tous les cas, la somme totale des probabilités est égale à l'unité; car tous les cas possibles sont ainsi décomposés en catégories exclusives l'une de l'autre.

Remarquons que, avant l'introduction des boules rouges dans l'urne, la probabilité d'amener une blanche et celle d'amener une noire étant dans le rapport des nombres a et b, ces probabilités demeurent encore dans le même rapport, quand on a introduit les boules rouges. Les nombres des cas favorables, en effet, n'ont pas changé; c'est seulement le nombre des cas possibles, qui s'est trouvé modifié par l'adjonction des rouges.

Considérons un disque divisé en n secteurs, colorés diversement, et dont les angles au centre soient proportionnels à n nombres positifs donnés $q_1, q_2, \ldots, q_n$. Si l'on fait tourner le disque devant un index, les différentes orientations étant supposées également probables et les coups où l'index tombe devant la ligne de séparation de deux secteurs limitrophes étant considérés comme des coups nuls, la probabilité que l'index soit au moment de l'arrêt en regard du secteur correspondant au nombre q_i sera

$$p_i = \frac{q_i}{q_1 + q_2 + \ldots + q_n}.$$

Les probabilités p_i seront donc n nombres *positifs de somme égale à l'unité* et l'on pourra, sauf cette seule et unique restriction, se les donner *arbitrairement*.

Jouons aux dés avec deux dés. On pourra amener tous les points compris entre as et as, soit 2 et 6 et 6, soit 12. Il est facile de former le tableau des probabilités de chaque point possible. On obtient ainsi :

Nombre de points........	2	3	4	5	6	7	8	9	10	11	12
Nombre de façons de les obtenir................	1	2	3	4	5	6	5	4	3	2	1
Probabilités..............	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Par exemple, on a 7 par les combinaisons $6+1$, $5+2$, $4+3$, $3+4$, $2+5$, $1+6$, c'est-à-dire de six façons distinctes. C'est le point le plus probable. Sa probabilité est $1/6^e$.

On suppose ici que les dés ne sont pas pipés, c'est-à-dire que, pour chacun d'eux et indépendamment de l'autre, chaque point de 1 à 6 est également probable.

Il y a lieu, dans chaque cas particulier, de bien préciser le mécanisme, suivant lequel on fait intervenir le hasard. Par exemple, je mets dans une urne neuf boules portant les numéros 0, 1, 2, 3, 4, 5, 6, 7, 8, ces boules étant identiques, si j'en tire une au hasard, la probabilité d'amener une quelconque de ces boules sera égale à $1/9^e$. Mais maintenant, je répartis ces 9 boules en 5 urnes

de la façon suivante :

Urne.	Nos.
1	0
2	1, 2
3	3, 4, 5
4	6, 7
5	8

Cela posé, admettons que tout d'abord je tire au sort le numéro de l'urne et qu'une fois l'urne trouvée, je tire la boule.

Je dis que les probabilités sont alors conformes au tableau suivant :

Nos.	Probabilité pour chacun des numéros.
0, 8	$\frac{1}{5}$
1, 2, 6, 7	$\frac{1}{10}$
3, 4, 5	$\frac{1}{15}$

La somme des huit probabilités est l'unité.

Pour justifier ces nombres, il suffit de remarquer que la probabilité pour chacun des nos 1, 2, 6, 7 est bien la moitié de celle des nos 0, 8 qui est évidente et égale à 1/5e, et celle des nos 3, 4, 5 n'est que le tiers de cette dernière.

Cet exemple prouve que la façon dont on fait jouer le hasard influe essentiellement sur les probabilités. Il faut donc bien dans chaque cas définir ce que les mots *au hasard* peuvent signifier. Cela revient à rechercher avec soin les conditions d'égale probabilité, quand elles n'ont pas été explicitées d'une façon précise.

En *résumé*, bien que d'après les principes de la mécanique les phénomènes observés soient toujours la conséquence rigoureuse de certains antécédents, conformément aux lois naturelles, il y a des événements dont nous ne pouvons prévoir l'issue à cause de l'imperfection de nos informations à l'égard de ces antécédents et parmi eux, il y en a pour lesquels différentes éventualités se présentent à l'esprit comme également probables. Des suppléments d'information peuvent d'ailleurs modifier nos prévisions primi-

tives. La notion de cas également probables est une conception abstraite, indépendante de toute réalisation, au même titre qu'en géométrie, celle de points alignés en ligne droite.

On aura donc soin de définir les cas que l'on considère comme également probables. La probabilité d'un événement est le rapport existant entre le nombre des cas favorables à l'arrivée de l'événement et le nombre total des divers cas possibles et également probables.

L'arrivée d'un événement et la non arrivée du dit sont des éventualités contraires ou contradictoires. Deux événements sont exclusifs l'un de l'autre si tout cas qui favorise l'un ne peut pas favoriser l'autre.

L'unité est le symbole de la certitude.

Le principe de la *probabilité totale* consiste en ce que, si deux événements ont les probabilités p_1 et p_2 et s'ils sont exclusifs l'un de l'autre, la probabilité de l'arrivée de l'un ou de l'autre est $p_1 + p_2$.

S'ils n'étaient pas exclusifs l'un de l'autre, il y aurait des cas favorables comptés deux fois, à savoir ceux qui sont favorables à l'un et à l'autre en même temps.

Le calcul des probabilités se réduit donc en définitive à deux opérations de dénombrement :

1° la décomposition de tous les cas possibles en cas également probables;

2° l'énumération des cas favorables à l'arrivée de l'événement attendu.

La probabilité de l'événement est le rapport du second nombre au premier. Et de cette façon, on voit que le calcul des probabilités s'appuie sur le calcul des permutations, arrangements et combinaisons. Par exemple, une urne contient q_1 blanches et q_2 noires, on tire d'un seul coup α boules. Quelle est la probabilité pour que ces α boules soient blanches?

Il faut évidemment tout d'abord que α soit au plus égal à q_1. Le nombre des boules étant $q_1 + q_2$, le nombre des façons d'en prendre α quelle que soit leur couleur est le nombre des combinaisons de $q_1 + q_2$ objets α à α, ou $C^{\alpha}_{q_1+q_2}$. C'est le nombre des cas également probables. Un cas favorable est celui où l'on tire α

blanches, or il y en a q_1 de cette couleur. On a donc $C_{q_1}^{\alpha}$ combinaisons possibles de q_1 blanches α à α. La probabilité cherchée est donc

$$C_{q_1}^{\alpha} : C_{q_1+q_2}^{\alpha}.$$

2. Probabilité composée. — Le calcul des probabilités repose essentiellement sur deux propositions fondamentales, l'une concerne la probabilité *composée*, l'autre la probabilité *totale*.

Afin de fixer les idées, je vais exposer tout d'abord la solution d'un problème correspondant à une application de la première proposition.

Une urne contient a blanches et b noires (il s'agit de boules), on tire une première boule, c'est un premier événement; sans la remettre dans l'urne, on tire une seconde boule, c'est un second événement. L'événement que l'on attend est un troisième événement qui résulte de la combinaison des deux, admettons qu'il consiste à amener une blanche au premier tirage et une noire au second.

On peut raisonner comme suit : le nombre total de cas également possibles consiste dans le produit du tirage d'une quelconque des $a+b$ boules au premier tour et du tirage d'une quelconque des $a+b-1$ boules au second tour. Le nombre des éventualités possibles (et également probables) est donc

$$(a+b)(a+b-1).$$

D'autre part, il s'agit d'évaluer le nombre des cas favorables à l'événement attendu, or il faut qu'au premier tour on ait tiré une blanche, le nombre des cas favorables est a, puis, en admettant qu'on ait bien amené une blanche au premier tour, ce qui ne change pas le nombre des noires restant dans l'urne, il faut qu'on amène ensuite une noire au second tour, le nombre des cas favorables est b. Le nombre des éventualités favorables distinctes est donc ab et c'est là le nombre des cas favorables à l'événement attendu, car les conditions envisagées comme nécessaires sont aussi suffisantes.

En conséquence la probabilité cherchée est

$$\frac{ab}{(a+b)(a+b-1)}.$$

Formons le tableau des éventualités possibles. Soient A le tirage qui amène une blanche a, B le tirage qui amène une noire b. Indiquons par le signe $+$ que l'événement A se produit, par le signe $-$ que c'est l'événement B. Il y aura quatre éventualités possibles suivant le tableau ci-après qui nous donne les résultats au premier et au second tirage et dans la troisième colonne, le nombre de cas correspondants.

1er tirage.	2e tirage.	Nombre de cas.
$+$	$-$	ab
$+$	$+$	$a(a-1)$
$-$	$-$	$b(b-1)$
$-$	$+$	ba

L'événement attendu résulte du concours de A au premier tirage et de B au second, ce qui correspond à la combinaison $(+-)$; nombre de cas ab. Le nombre de cas possibles est

$$ab + a(a-1) + b(b-1) + ba = (a+b)(a+b-1).$$

Donc la probabilité P est

$$P = \frac{ab}{(a+b)(a+b-1)}.$$

Pour tirer de cette expression la forme qui correspond au théorème de la probabilité composée nous posons

$$p = \frac{a}{a+b}, \qquad q' = \frac{b}{a+b-1}$$

et nous avons

$$P = p.q',$$

relation qu'il s'agit d'interpréter :

Qu'est-ce que p?

C'est la probabilité d'amener une blanche au premier coup. Le tableau permet de vérifier ce résultat d'ailleurs évident, car le nombre de cas où nous trouvons le signe $+$ au premier tirage est $ab + a(a-1)$ ou à $a(a+b-1)$ et le nombre total de cas possibles est $(a+b)(a+b-1)$, comme nous l'avons vu; le rapport de ces deux nombres est bien $\frac{a}{a+b}$, c'est-à-dire p.

Qu'est-ce que p'?

Après qu'on a tiré une blanche au premier tirage, on sait désormais que le nombre de boules qui restent dans l'urne est composé de $a - 1 + b$ boules dont b sont noires. Donc q' est la probabilité de tirer une noire au second tour, quand une blanche a été tirée de l'urne au premier.

C'est ce que le tableau permet encore de vérifier, car si l'on sait qu'une blanche a été tirée au premier coup, les combinaisons à considérer sont $(+-)$ et $(++)$, qui se peuvent présenter respectivement dans des nombres de cas égaux à ab et $a(a+1)$, la probabilité q' est donc bien égale à

$$\frac{ab}{ab + a(a-1)} = \frac{b}{a+b-1},$$

comme nous l'avons trouvé directement.

En conséquence, la probabilité P est égale au produit de la probabilité p de tirer une blanche au premier coup, par la probabilité que nous attribuons au tirage d'une noire au second coup, quand nous savons qu'au premier, c'est une blanche qui a été tirée.

Donnons à cette démonstration une tournure générale. Soient E et E' deux événements et considérons l'événement qui résulte du concours de E et E'. Il y a quatre hypothèses à envisager : les événements E et E' peuvent se produire tous les deux ou ils peuvent ne se produire ni l'un ni l'autre, enfin un seul des deux E ou E' peut se produire. Marquons par le signe + le cas où l'événement considéré se produit et par le signe — celui où il ne se produit pas. Nous pouvons former le tableau.

E.	E'.	Nombre de cas.
+	+	α
+	—	β
—	+	γ
—	—	δ

La probabilité cherchée P est

$$P = \frac{\alpha}{\alpha+\beta+\gamma+\delta}.$$

La probabilité p pour que E se produise est

$$p = \frac{\alpha+\beta}{\alpha+\beta+\gamma+\delta}.$$

La probabilité q' pour que E′ se produise quand on sait que E s'est produit est

$$q' = \frac{\alpha}{\alpha + \beta}.$$

D'où

$$P = pq'.$$

On peut aussi dire : en intervertissant les rôles de E et de E′. La probabilité pour que (E E′) se produise est toujours

$$P = \frac{\alpha}{\alpha + \beta + \gamma + \delta}.$$

La probabilité q pour que E′ se produise, soit seul, soit avec E est

$$q = \frac{\alpha + \gamma}{\alpha + \beta + \gamma + \delta}.$$

La probabilité p' pour que E se produise, quand on sait que E′ s'est produit est

$$p' = \frac{\alpha}{\alpha + \gamma}.$$

D'où

$$P = qp'.$$

Donc

$$pq' = qp'.$$

3. **Événement indépendant.** — Au point de vue du calcul des probabilités l'événement E est indépendant de l'événement E′ si la connaissance que l'événement E′ s'est produit n'influe pas sur la probabilité à attribuer à l'événement E. On a alors

$$p' = p$$

et il en résulte nécessairement

$$q' = q.$$

Donc si E est indépendant de E′, réciproquement E′ est indépendant de E.

Exemple. — Une urne contient les numéros successifs de 1 à 12 inclusivement.

Quelle est la probabilité d'amener un point pair? Réponse $\frac{1}{2}$.

Mais, après que le numéro a été tiré de l'urne, et avant que nous

ne l'ayons vu, on nous dit que le point tiré est un multiple de 3, la probabilité est-elle changée? Réponse : non.

Pourquoi? parce que les numéros multiples de 3 sont 3, 6, 9, 12 et que la probabilité d'un point pair est encore $\frac{1}{2}$.

Conséquence. — Ces deux événements, le point est pair, le point est multiple de 3, sont indépendants. Appliquons le théorème de la probabilité composée.

Quelle est la probabilité d'amener un point multiple de 2 et de 3, toujours de la même urne?

La probabilité d'un point pair est $\frac{1}{2}$.

La probabilité d'un point impair est $\frac{1}{3}$.

Ces deux éventualités étant indépendantes, la probabilité d'amener un point multiple à la fois de 2 et de 3 est $\frac{1}{6}$; et effectivement 12 et 6 sont les seuls numéros favorables sur les 12 contenus dans l'urne.

Changeons la constitution de l'urne, en supprimant le numéro 12. Posons le même problème :

Quelle est la probabilité d'amener un point multiple à la fois de 2 et de 3? Réponse : $\frac{1}{11}$. Or on a les probabilités suivantes.

La probabilité p d'amener un point pair est $\frac{5}{11}$.

La probabilité q d'amener un point multiple 3 est $\frac{3}{11}$.

Mais comme les deux événements ne sont pas indépendants, on ne peut répondre à la question en faisant le produit.

Cependant, les points pairs étant 2, 4, 6, 8, 10, on voit que la probabilité q' d'en tirer un numéro multiple de 3 est $\frac{1}{5}$, correspondant au cas favorable où l'on tire 6; et l'on a :

$$p = \frac{5}{11}, \qquad q' = \frac{1}{5}.$$

D'où

$$P = pq' = \frac{5}{11} \times \frac{1}{5} = \frac{1}{11},$$

ce qui est la solution correcte.

Le nombre 6 est le seul qui sur les onze soit à la fois multiple à la fois de 2 et de 3.

Remarque. — Étant donnée une urne qui contient a boules blanches et b boules noires, la règle $pq' = pq'$ montre que les deux problèmes suivants sont équivalents :

1° Quelle est la probabilité pour que, tirant de l'urne successivement deux boules, l'une soit blanche, l'autre noire, en admettant qu'on ne remette pas dans l'urne la première boule tirée?

2° Quelle est, dans les mêmes conditions, la probabilité d'avoir une noire, puis une blanche?

Les deux problèmes ne diffèrent que par l'ordre dans lequel on doit avoir deux boules de couleurs différentes.

4. Probabilité totale. — Afin de fixer les idées, commençons par traiter un problème particulier.

Une urne contient les numéros 1, 2, 3, 4, 5. On en tire deux d'un coup. Quelle est la probabilité, pour que sur les deux numéros tirés il y en ait au moins un qui soit pair?

Formons le tableau des cas possibles

12	13	14	15
	23	24	25
		34	35
			45

Sur les 10 cas possibles, il y en a 7 favorables; la réponse est donc

$$P = \frac{1}{7}.$$

mais on peut dire autrement.

La combinaison favorable contiendra soit 2, soit 4.

Or chacun des nombres 2 et 4 peut être associé avec chacun des quatre autres, d'où 8 combinaisons; il y a à retrancher une unité à ce nombre, parce que la combinaison (2, 4) a été comptée deux fois au lieu d'une, comme appartenant aux deux catégories, issues respectivement de 2 et de 4. On retrouve donc le chiffre correct $\frac{1}{7}$.

D'une façon générale, nous énoncerons le théorème suivant :

THÉORÈME. — *Si un événement peut se produire soit par l'arrivée d'un premier événement* E, *soit par l'arrivée d'un second événement* E′, *sa probabilité est égale à la somme des probabilités de* E *et de* E′ *respectivement, diminuée de la probabilité pour que* E *et* E′ *se produisent ensemble.*

Reprenons le tableau où E_0 et E'_0 sont les événements contraires de E et de E′ :

EE′........................	α
EE'_0........................	β
E_0E'........................	γ
$E_0E'_0$........................	δ

L'événement attendu se produisant soit par l'arrivée de E, soit par l'arrivée de E′, le nombre de cas qui lui sont favorables est $\alpha + \beta + \gamma$ et la probabilité cherchée x est

$$x = \frac{\alpha + \beta + \gamma}{\alpha + \beta + \gamma + \delta}.$$

Or, la probabilité de E est

$$\frac{\alpha + \beta}{\alpha + \beta + \gamma + \delta}.$$

Celle de E′ est

$$\frac{\alpha + \gamma}{\alpha + \beta + \gamma + \delta},$$

et l'on a identiquement

$$x = \frac{\alpha + \beta}{\alpha + \beta + \gamma + \delta} + \frac{\alpha + \gamma}{\alpha + \beta + \gamma + \delta} - \frac{\alpha}{\alpha + \beta + \gamma + \delta},$$

ce qui est l'expression analytique du théorème, puisque la fraction à soustraire représente la probabilité de l'arrivée de E et de E′ ensemble.

Nota. — Si les événements E et E′ sont exclusifs l'un de l'autre, ils ne peuvent se produire simultanément et l'on a $\alpha = 0$. En ce cas, la probabilité de l'événement attendu est la somme des probabilités relatives à l'événement E et à l'événement E′.

5. **Principe de Bayes.** — Avant d'énoncer ce principe, recourons

encore à un exemple qui puisse lui servir d'introduction. Considérons le jet de deux dés et formons le tableau :

Nombre de points amenés.	2	3	4	5	6	7	8	9	10	11	12
Probabilité correspondante.	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Comparons les probabilités des deux propositions suivantes :

1° le point est pair; soit p_1 la probabilité correspondante;

2° le point est multiple de 3; soit p_2 la probabilité correspondante.

On a

$$p_1 = \frac{1+3+5+5+3+1}{36} = \frac{18}{36} = \frac{1}{2},$$

$$p_2 = \frac{2+5+4+1}{36} = \frac{12}{36} = \frac{1}{3}.$$

Nous avons tenu compte de toutes les façons qu'on a d'amener un point, mais comme notre intention est de ne comparer que la probabilité d'amener un point pair avec celle d'amener un point multiple de trois, les autres éventualités nous sont indifférentes, je veux dire celles où le point amené n'est divisible ni par 2, ni par 3, supprimons-les. On a alors le tableau suivant des cas possibles et du nombre de façons d'obtenir chacun d'eux.

Nombre de points amenés...	2	3	4	6	8	9	10	12
Nombre de manières.......	1	2	3	5	5	4	3	1

Le nombre des cas conservés est de 24. Par conséquent, les probabilités sont les suivantes :

Probabilités...............	$\frac{1}{24}$	$\frac{2}{24}$	$\frac{3}{24}$	$\frac{5}{24}$	$\frac{5}{24}$	$\frac{4}{24}$	$\frac{3}{24}$	$\frac{1}{24}$

Mais ces nouvelles probabilités sont aux précédentes dans un rapport constant celui de $\frac{1}{24}$ à $\frac{1}{36}$. Le rapport des probabilités relatives aux deux propositions énoncées n'a donc pu changer et l'on a dans les deux cas, en appelant p'_1 et p'_2 les nouvelles probabilités

$$\frac{p_1}{p_2} = \frac{\frac{1}{2}}{\frac{1}{3}} = \frac{3}{2} = \frac{p'_1}{p'_2}.$$

On aurait avec le second tableau

$$p'_1 = \frac{18}{24}, \qquad p'_2 = \frac{12}{24} \qquad \text{ou} \qquad p'_1 = \frac{3}{4}, \qquad p'_2 = \frac{1}{2}.$$

On voit donc que p'_1 et p'_2 sont parfaitement distincts de p_1 et de p_2, mais que cela n'empêche pas que leur rapport demeure inchangé. Les cas favorables à l'une ou à l'autre éventualité sont restés les mêmes, le nombre total de cas considérés a seul été modifié.

Voici maintenant le principe de Bayes.

Lorsqu'un renseignement nouveau élimine certains cas considérés d'abord comme possibles, les probabilités des cas qui restent possibles prennent des valeurs nouvelles, mais ces valeurs sont proportionnelles aux anciennes.

Si leur somme doit être égale à l'unité, c'est-à-dire si les éventualités conservées sont celles d'événements exclusifs, dont l'ensemble représente le tableau de tous les cas possibles, la règle détermine alors complètement les nouvelles probabilités.

Revenons à l'exemple concernant le jet de deux dés et les probabilités relatives d'un point pair ou d'un point multiple de 3 et ce qui arrive quand on supprime les points qui ne sont ni pairs, ni multiples de 3.

Dans cet exemple, le renseignement nouveau consistait à admettre que les points 5 et 7 n'avaient pas à être pris en considération. Si l'on détermine tous les cas où ces points auraient été amenés, il ne reste plus en ligne de compte que 24 cas possibles.

Or on avait

$$\frac{p_1}{p_2} = \frac{3}{2},$$

nous aurons donc

$$\frac{p'_1}{p'_1} = \frac{3}{2}.$$

Mais pouvons-nous écrire que la somme est l'unité ? savoir

$$p'_1 + p'_2 = 1.$$

Non, car un point peut être à la fois multiple de 2 et de 3, c'est le cas de 6 ou de 12.

La probabilité d'avoir 6 ou 12 est

$$\frac{5+1}{24} = \frac{6}{24} = \frac{1}{4}.$$

Appelons-la $p'_{1,2}$; comme on a identiquement

$$\frac{18}{24} + \frac{12}{24} - \frac{6}{24} = \frac{24}{24} = 1,$$

on pourra écrire

$$p'_1 + p'_2 - p'_{1,2} = 1,$$

ce qui est conforme à la règle relative à la probabilité totale. Celle-ci est égale à la somme des probabilités d'avoir un point multiple de 2 ou un point multiple de 3, diminuée de la probabilité d'avoir un point multiple, à la fois, de 2 et de 3.

5. **Application du principe de Bayes.** — Une collection d'urnes comprend une proportion p_1 d'urnes U_1 identiques, contenant une proportion q_1 de boules blanches, une proportion p_2 d'urnes U_2 identiques entre elles, contenant une proportion q_2 de boules blanches, et d'autres urnes encore, mais celles-ci ne contenant pas de boules blanches. On nous fait savoir qu'une boule blanche a été tirée et l'on demande la probabilité P_1 qu'elle ait été tirée de l'urne U_1 ainsi que la probabilité P_2 qu'elle ait été tirée de l'urne U_2.

Il est évident que, les autres urnes ne contenant pas de blanches et les deux éventualités considérées étant exclusives, on a ici

$$P_1 + P_2 = 1.$$

Il suffit donc de déterminer le rapport de P_1 à P_2.

Or on nous dit qu'une blanche a été tirée; écartons cette information et disons simplement une boule a été tirée; la probabilité pour que l'on obtienne une blanche de l'urne U_1 est d'après le principe de la probabilité composée égale à $p_1 q_1$, puisque p_1 est la probabilité de choisir l'urne U_1 et q_1 la probabilité d'amener une blanche quand U_1 a été choisi; et de même la probabilité pour qu'on obtienne une blanche de l'urne U_2 est égale à $p_2 q_2$. Il y a bien d'autres événements à considérer, mais leurs probabilités respectives ne nous intéressent pas, car elles ne se rapportent pas

au tirage d'une blanche. Je pose

$$\alpha_1 = p_1 q_1, \qquad \alpha_2 = p_2 q_2,$$

α_1 est la probabilité de tirer une blanche de l'urne U_1 ;
α_2 est la probabilité de tirer une blanche de l'urne U_2.

Mais ces probabilités se rapportent à l'hypothèse fondamentale, que nous ne savons encore rien sur l'issue du tirage. Ce sont des probabilités *a priori*. Maintenant, nous écartons tous les cas où ce n'est pas une blanche qui a été tirée, les probabilités que l'on ait tiré la blanche de l'urne U_1 ou de l'urne U_2 ne restent plus égales à α_1 et à α_2, elles prennent de nouvelles valeurs, celles que nous avons appelées P_1 et P_2, mais, en vertu du principe de Bayes, on doit avoir

$$\frac{P_1}{\alpha_1} = \frac{P_2}{\alpha_2}.$$

D'où l'on déduit

$$\frac{P_1}{\alpha_1} = \frac{P_2}{\alpha_2} = \frac{P_1 + P_2}{\alpha_1 + \alpha_2} = \frac{1}{\alpha_1 + \alpha_2},$$

c'est-à-dire

$$P_1 = \frac{\alpha_1}{\alpha_1 + \alpha_2}, \qquad P_2 = \frac{\alpha_2}{\alpha_1 + \alpha_2}.$$

ou encore

$$P_1 = \frac{p_1 q_1}{p_1 q_1 + p_2 q_2}, \qquad P_2 = \frac{p_2 q_2}{p_1 q_1 + p_2 q_2}.$$

On peut considérer le choix de l'urne U_1 ou de l'urne U_2 respectivement, comme celui de deux causes ayant l'une ou l'autre pour conséquence le tirage de la boule blanche considérée et l'on dit alors, comme cela arrive souvent dans la littérature, que le problème résolu est celui de la *probabilité des causes*.

On interprétera ces formules dans le langage de la statique, en imaginant qu'on porte sur l'axe Ox aux points d'abscisses q_1 et q_2 des masses égales respectivement à p_1 et à p_2.

Si alors, j'appelle x la valeur commune des deux rapports

$$\frac{p_1 q_1}{P_1} = \frac{p_2 q_2}{P_2} = x.$$

j'aurai

$$x P_1 = p_1 q_1, \qquad x P_2 = p_2 q_2, \qquad P_1 + P_2 = 1$$

et je pourrai dire que x est la distance à laquelle il faut placer une

masse égale à l'unité pour que son moment par rapport à l'origine soit égal au moment résultant des masses p_1 et p_2.

Cela résulte de ce que j'ai

$$(P_1 + P_2)x = p_1 q_1 + p_2 q_2.$$

P_1 et P_2 sont les portions de la masse unité qui, placées à la distance x ont respectivement même moment que p'_1 et p_2.

Voici encore quelques explications détaillées sur l'application du principe de Bayes.

Considérons une urne contenant n_1 boules blanches, n_2 boules noires et n_3 boules de couleur et qu'on se propose d'en extraire une au hasard, en supposant que l'éventualité de tirer une quelconque des boules de l'urne a toujours la même probabilité.

Alors les probabilités p_1, p_2, p_3 de tirer une blanche, une noire ou une boule de couleur sont, en appelant N le nombre total des boules,

$$p_1 = \frac{n_1}{N}, \qquad p_2 = \frac{n_2}{N}, \qquad p_3 = \frac{n_3}{N}, \qquad n_1 + n_2 + n_3 = N.$$

Si maintenant on remplace n_3 par n'_3, les nouvelles probabilités, le nombre total des boules étant devenu N', seront

$$p'_1 = \frac{n_1}{N'}, \qquad p'_2 = \frac{n_2}{N'}, \qquad p'_3 = \frac{n'_3}{N'}, \qquad n_1 + n_2 + n'_3 = N$$

et l'on aura

$$\frac{p_1}{p'_1} = \frac{p_2}{p'_2}.$$

De plus, si l'on avait

$$n'_3 = 0,$$

on aurait

$$p'_1 + p'_2 = 1,$$

puisque l'on épuise tous les cas possibles en considérant le tirage d'une blanche ou d'une noire.

Cherchons maintenant à appliquer le principe de Bayes au cas où l'on supprimerait les n_3 boules de couleur et *une* boule blanche.

Je commencerai par remplacer une boule blanche par une boule de couleur, ce qui remplacera le nombre n_1 par le nombre $n_1 - 1$ et le nombre n_3 par le nombre $n_3 + 1$ et les probabilités initiales



dont je partirai P_1, P_2, P_3 seront

$$P_1 = p_1 - \frac{1}{N}, \qquad P_2 = p_2, \qquad P_3 = p_3 + \frac{1}{N}.$$

Si je supprime alors les $n_3 + 1$ boules de couleur, je pourrai appliquer le principe de Bayes et j'aurai les nouvelles probabilités P'_1, P'_2 et zéro, avec

$$\frac{P'_1}{P_1} = \frac{P'_2}{P_2} \quad \text{et} \quad P'_1 + P'_2 = 1.$$

D'où

$$P'_1 = \frac{P_1}{P_1 + P_2}, \qquad P'_2 = \frac{P_2}{P_1 + P_2}.$$

En résumé, dans l'application du principe de Bayes, il faut que les cas supprimés soient des cas complémentaires de ceux que l'on conserve.

6. **Application du calcul des probabilités à l'analyse combinatoire.** — 1° Étant données m boules numérotées de 1 à m, d'ailleurs identiques, placées dans une urne. On en tire n et l'on demande la probabilité pour que leurs numéros soient ceux de n boules désignées à l'avance.

Tirons d'abord une boule, il y a n façons d'amener une boule de la combinaison sur les m façons totales différentes d'amener une boule. La probabilité est donc $\frac{n}{m}$. Mais quand on a tiré une boule appartenant à la combinaison, boule qu'on se garde de remettre dans l'urne, il n'y a plus que $m - 1$ boules en tout et $n - 1$ d'entre elles appartiennent à la combinaison, donc quand on tire une seconde boule, la probabilité pour qu'elle soit de la combinaison est $\frac{n-1}{m-1}$. On a donc pour probabilité d'amener deux boules appartenant à la combinaison le produit de ces deux probabilités soit $\frac{n(n-1)}{m(m-1)}$. En continuant à tirer une boule à chaque fois jusqu'à la $n^{\text{ième}}$ épreuve, et répétant le même raisonnement, on voit que la probabilité que les n boules que l'on tire de l'urne soient celles qui constituent la combinaison considérée est

$$p = \frac{n(n-1)\ldots 2.1}{m(m-1)\ldots(m-n+1)}.$$

Il en résulte évidemment que le nombre C_m^n de combinaisons distinctes est l'inverse $\frac{1}{p}\cdot$ Car si l'on fait la somme des probabilités, d'ailleurs toutes égales entre elles, relatives à toutes les C_m^n combinaisons possibles, on a la probabilité totale qui est manifestement égale à l'unité, puisque, après avoir tiré les n boules, on a forcément tiré l'une des C_m^n combinaisons distinctes. On a donc

$$C_m^n = \frac{m(m-1)\ldots(m-n+1)}{1.2\ldots n}.$$

A toute combinaison de m chiffres n à n, correspond une combinaison et une seule des chiffres exclus de la combinaison, on a donc

$$C_m^n = C_m^{m-n}.$$

On peut d'ailleurs écrire

$$C_m^n = C_m^{m-n} = \frac{m!}{n!\,(m-n)!}$$

ou encore, en posant

$$n = \alpha, \qquad m - n = \beta,$$

$$C_m^n = f(\alpha, \beta) = \frac{(\alpha+\beta)!}{\alpha!\,\beta!}.$$

2° Étant données m boules numérotées de 1 à m, mais d'ailleurs identiques, placées dans une urne, on les tire successivement toutes de l'urne et l'on demande la probabilité de les tirer dans un ordre déterminé.

La première boule est donc une boule donnée, la probabilité de la tirer est $\frac{1}{m}\cdot$ Dès lors, c'est tout de suite le même problème qui se pose, mais pour $m-1$ boules. D'après le théorème de la propabilité composée, on a donc pour la probabilité cherchée

$$p = \frac{1}{m(m-1)\ldots 2.1} \qquad \text{ou} \qquad p = \frac{1}{m!}.$$

Si donc j'appelle P_m le nombre des permutations distinctes, j'aurai

$$P_m p = 1 \qquad \text{ou} \qquad P_m = m!$$

$P_m p$ est la probabilité totale qui est ici la certitude, car si

l'on tire m boules successivement, on a certainement l'une des permutations et il y a P_m façons d'opérer ainsi avec la probabilité p.

On peut écrire

$$C_m^n = \frac{P_m}{P_n P_{m-n}}.$$

3° Considérons maintenant un papier quadrillé ou, si l'on veut, la portion du plan comprise dans l'angle droit xOy, après l'avoir décomposé en carrés contigus de côté égal à l'unité par des parallèles aux axes, menés par les points de ces axes dont les cotes sont des nombres entiers. J'appelle (1, 1) le carré placé dans l'angle à l'origine des axes et (α, β) le carré de rang α dans la ligne de rang β, c'est-à-dire le carré commun à la $\alpha^{\text{ième}}$ colonne et à la $\beta^{\text{ième}}$ ligne.

Sur cette sorte d'échiquier, une pièce mobile est censée ne pouvoir se déplacer au cours de son mouvement que d'un pas égal à un dans le sens Ox ou d'un pas égal à un dans le sens Oy.

On la déplace suivant cette règle de la case (1, 1) jusqu'à la case (α, β) et l'on demande combien il y a de chemins distincts pour obtenir ce résultat.

Soit $f(\alpha, \beta)$ ce nombre. On a évidemment

$$f(\alpha, 1) = f(1, \beta) = 1,$$

parce qu'il n'y a qu'une seule façon de se déplacer sur les bandes qui forment la première colonne ou la première ligne. D'autre part, on a la formule de récurrence

$$f(\alpha + 1, \beta + 1) = f(\alpha, \beta + 1) + f(\alpha + 1, \beta),$$

parce qu'en arrivant sur la case $(\alpha + 1, \beta + 1)$ la pièce, pour faire son dernier pas, a dû nécessairement partir de la case précédente située dans la même ligne ou dans la même colonne.

On peut donc évaluer de proche en proche la valeur de $f(\alpha, \beta)$ et en inscrivant sur chaque case le nombre de chemins obtenus, on forme le tableau suivant qui est la clé du problème

O → x, ↓ y

1	1	1	1	1
1	2	3	4	
1	3	6		
1	4			
1				

On a, d'ailleurs,

$$\begin{aligned}
f(\alpha+1, \beta+1) &= f(\alpha+1, \beta) + f(\alpha, \beta+1),\\
f(\alpha+1, \beta) &= f(\alpha+1, \beta-1) + f(\alpha, \beta),\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
f(\alpha+1, 2) &= f(\alpha+1, 1) + f(\alpha, 2),\\
f(\alpha+1, 1) &= f(\alpha, 1).
\end{aligned}$$

D'où

$$f(\alpha+1, \beta+1) = f(\alpha, 1) + f(\alpha, 2) + \ldots + f(\alpha, \beta) + f(\alpha, \beta+1),$$

ou bien

$$f(\alpha+1, \beta+1) = \sum_{i=0}^{\beta} f(\alpha, i+1).$$

(Rappelons, pour l'application des équations qui vont suivre, comment on effectue une somme arithmétique S de la forme suivante :

$$S = 1.2.3 + 2.3.4 + 3.4.5 + 4.5.6 + 5.6.7.$$

Multiplions chaque terme par 4 et effectuons les sommes successives :

$$\begin{aligned}
1.2.3.4 + 2.3.4.4 &= 2.3.4.5,\\
2.3.4.5 + 3.4.5.4 &= 3.4.5.6,
\end{aligned}$$

et ainsi de suite, il vient

$$4S = 5.6.7.8).$$

Nous avons alors successivement :

$$\begin{aligned}
f(2, \beta+1) &= \sum_{i=0}^{\beta} f(1, i+1) = \beta+1,\\
f(3, \beta+1) &= \sum f(2, i+1) = \frac{(\beta+1)(\beta+2)}{1.2},\\
f(4, \beta+1) &= \sum f(3, i+1) = \frac{(\beta+1)(\beta+2)(\beta+3)}{1.2.3},\\
&\ldots\ldots\ldots\ldots\ldots\ldots,\\
f(\alpha+1, \beta+1) &= \frac{(\beta+1)(\beta+2)\ldots(\beta+\alpha)}{1.2\ldots\alpha}.
\end{aligned}$$

On a donc

$$f(\alpha+1, \beta+1) = \frac{(\alpha+\beta)!}{\alpha!\,\beta!} = C_{\alpha+\beta}^{\alpha} = C_{\alpha+\beta}^{\beta}.$$

Le nombre de chemins distincts permettant de déplacer la pièce mobile de la case (1 . 1) à la case ($\alpha + 1$, $\beta + 1$) est égal, d'après la formule trouvée, au nombre de combinaisons α à α de $\alpha + \beta$ objets. Mais, si nous montrons que le nombre cherché doit *a priori* représenter ce nombre de combinaisons, nous aurons par là même trouvé une autre méthode pour obtenir la formule donnant C_m^n.

Or pour aller de la case (1, 1) à la case ($\alpha + 1$, $\beta + 1$), il faut faire α pas dans le sens Ox et β pas dans le sens Oy, peu importe d'ailleurs l'ordre dans lequel ces pas ont été exécutés.

Numérotons les pas de 1 à $\alpha + \beta$. Il y en a α dirigés suivant Ox et β dirigés suivant Oy. Donc, pour chaque chemin, on prend dans la série des numéros de 1 à $\alpha + \beta$ un nombre α de numéros, c'est-à-dire une combinaison parmi toutes celles de $\alpha + \beta$ numéros α à α. Et la correspondance est réciproque et univoque. Le nombre des chemins est donc égal à $C_{\alpha+\beta}^{\alpha}$ ou $C_{\alpha+\beta}^{\beta}$, les deux nombres devant être forcément égaux, car à chaque combinaison que l'on prend, correspond la combinaison complémentaire qu'on laisse.

4° Considérons maintenant le mécanisme suivant, par lequel nous allons faire intervenir le hasard pour déterminer le chemin que la pièce considérée devra suivre sur le damier. Imaginons une urne qui contienne des boules blanches et noires dans les proportions p et q respectivement, de façon qu'on puisse prendre $p + q$ égal à l'unité. Si p et q étaient incommensurables, il serait loisible d'imaginer, au lieu de l'urne, un disque divisé en deux secteurs de couleurs différentes, dont les surfaces seraient dans le rapport de p à q; on le ferait tourner devant un index. Supposons donc que, pour déterminer si le pas de la pièce sera dirigé suivant Ox ou suivant Oy, on procède au tirage d'une boule; si elle est blanche, le pas sera fait suivant Ox, si elle est noire, il sera suivant Oy. Puis on remet la boule dans l'urne pour conserver à celle-ci une composition constante. Alors la probabilité de faire parcourir à la pièce un certain chemin pour l'amener de la case initiale (1,1) à la case finale ($\alpha + 1$, $\beta + 1$) par α pas suivant Ox et β pas suivant Oy, dans un ordre déterminé sera égal, en vertu du théorème de la probabilité composée à la valeur d'un produit des facteurs p et q en nombre α pour p et en nombre β pour q, chacun à un rang bien déterminé; et chacun des produits ainsi

obtenus avec α facteurs p et β facteurs q aura la même valeur $(p^\alpha q^\beta)$, puisqu'un produit est indépendant de l'ordre des facteurs. On a donc des probabilités égales. Alors, d'après le principe de la probabilité totale, pour avoir la probabilité relative au déplacement de la pièce de la case initiale à la case finale considérée, quand on fait abstraction du tracé du chemin, il n'y aura qu'à multiplier par $C_{\alpha+\beta}^\alpha$ la probabilité $p^\alpha q^\beta$ d'un chemin particulier. Donc la probabilité est

$$C_{\alpha+\beta}^\alpha p^\alpha q^\beta.$$

Et nous pouvons remarquer aussitôt que si l'on considère toutes les cases qui correspondent à un nombre total $m = \alpha + \beta$ déterminé, il suffira de faire varier α de 1 à m; on aura ainsi toutes les probabilités des cas possibles; on aura donc la certitude et l'on doit avoir

$$1 = (p+q)^m = \sum_\alpha C_m^\alpha p^\alpha q^{m-\alpha}.$$

On reconnait la formule connue du binome de Newton.

SUR LA RÉPÉTITION DES ÉPREUVES ET LE PRINCIPE DE BAYES.

Considérons un jeu, où, à chaque coup, l'événement E a la probabilité p et son contraire a la probabilité q, de sorte qu'on a

$$p + q = 1.$$

On répète m fois l'épreuve. L'événement E peut se présenter α fois à des rangs déterminés et son contraire aux β autres fois, en posant comme plus haut

$$\alpha + \beta = m.$$

La probabilité de l'apparition de l'événement E à α rangs prédéterminés est un événement composé $\mathcal{E}_\alpha$ dont la probabilité $\mathcal{P}_\alpha$ est

$$\mathcal{P}_\alpha = p^\alpha q^\beta.$$

Mais il y a C_m^α événements analogues qui ont même probabilité et qui ne diffèrent entre eux que par le rang des α apparitions de l'événement E. J'appelle E_α l'événement composé qui résulte de

l'apparition de l'événement E un nombre de fois égal à α, quand on n'a pas égard aux rangs de ces apparitions et soit P_α la probabilité de E_α. J'ai, comme nous le savons :

$$P_\alpha = C_m^\alpha p^\alpha q^\beta.$$

Le nombre total de cas, quand on a égard aux rangs est 2^m et l'on a effectivement :

$$\sum_{\alpha=0}^{m} C_m^\alpha = (1+1)^m = 2^m.$$

Je suppose maintenant que sur les C_m^α cas où l'événement E_α se produit, on n'en conserve qu'un seul comme échantillon de cette catégorie et que l'on considère comme nuls tous les cas où E_α se produit d'une autre manière. Faisons de même pour toutes les valeurs de α.

Remarquons que comme nous ne conservons qu'un échantillon dans chaque catégorie de cas également possibles, il est loisible, sans changer les probabilités, de supposer par exemple que l'on a conservé comme échantillon de E_α l'arrangement A_α dans lequel l'événement E s'est produit aux α premiers coups.

Admettons ces diverses hypothèses, alors on n'a plus que $m+1$ cas possibles $A_0 A_1 \ldots A_m$ mais la probabilité *a priori* de A_α avant toutes ces suppressions était $\mathcal{P}_\alpha$. Si j'appelle Π_α la nouvelle probabilité de A_α, les nouvelles valeurs Π_α seront proportionnelles aux anciennes $\mathcal{P}_\alpha$ et l'on aura

$$\frac{\Pi_0}{q^m} = \frac{\Pi_1}{pq^{m-1}} = \ldots = \frac{\Pi_\alpha}{p^\alpha q^\beta} = \ldots = \frac{\Pi_m}{p^m}.$$

Cette suite de rapports égaux est égale à la fraction qu'on obtient, en divisant la somme des numérateurs, qui doit être égale à l'unité, par la somme des dénominateurs qui est

$$\frac{p^{m+1} - q^{m+1}}{p - q}.$$

On a donc la probabilité cherchée Π_α par la formule

$$\Pi_\alpha = \frac{p-q}{p^{m+1} - q^{m+1}} p^\alpha q^\beta \qquad (\alpha + \beta = m).$$

Et ce résultat peut aussi être obtenu par le raisonnement direct, sans recourir, comme nous l'avons fait, au principe de Bayes. Il suffit de remarquer que le cas $\mathcal{E}_\alpha$ diffère du cas $\mathcal{E}_{\alpha-1}$ en ce que, au $\alpha^{\text{ième}}$ coup, c'est la probabilité p qui a joué au lieu de la probabilité q. On a donc

$$\frac{\Pi_\alpha}{p} = \frac{\Pi_{\alpha-1}}{q}, \dots \frac{\Pi_1}{p} = \frac{\Pi_0}{q}.$$

Le principe de la probabilité totale donne

$$\sum_{\alpha=0}^{m} \Pi_\alpha = 1,$$

ce qui détermine Π_0 et redonne la formule précédente.

7. Épreuves répétées (cas des grands nombres). — Je me propose de faire une étude détaillée de la probabilité $P(\alpha)$

$$P(\alpha) = \frac{(\alpha+\beta)!}{\alpha!\,\beta!} p^\alpha q^\beta.$$

Et tout d'abord, je cherche, en considérant le nombre des épreuves $(\alpha+\beta)$ comme une constante donnée quelle est la valeur de α pour laquelle $P(\alpha)$ atteint son maximum.

Envisageons un cas particulier : si p et q sont égaux, et par conséquent tous les deux égaux à $\frac{1}{2}$, nous voyons que nous avons affaire aux termes successifs du binome de Newton divisés par 2^m, en posant $\alpha+\beta=m$. Et il est clair qu'alors les termes qui ont la plus grande valeur sont les deux termes du milieu, si m est impair, et le terme du milieu si m est pair.

Si p et q sont inégaux, je dis que le terme qui a la plus grande valeur est celui pour lequel α prend la valeur α_0, qui est le plus grand entier contenu dans $p(m+1)$.

Si, en effet, le terme qui a la plus grande valeur est $P(\alpha)$, on doit avoir

$$P(\alpha) > P(\alpha-1) \quad \text{et} \quad P(\alpha) > P(\alpha+1),$$

ce qui donne

$$\frac{p}{\alpha} > \frac{q}{\beta+1} \quad \text{et} \quad \frac{q}{\beta} > \frac{p}{\alpha+1}.$$

Remplaçons β par $m-\alpha$ et tenons compte de $p+q=1$, il reste pour déterminer la valeur de α, qui est un entier,

$$p(m+1)>\alpha>p(m+1)-1;$$

c'est cette valeur que j'appelle α_0. La valeur correspondante de β est β_0 et l'on a alors

$$q(m+1)>\beta>q(m+1)-1.$$

La proposition est donc démontrée.

D'autre part, comme α_0 et pm sont tous deux inférieurs à $p(m+1)$ et en diffèrent de moins d'une unité, ils ont entre eux une différence, qui, en valeur absolue, est inférieure à 1. On peut donc écrire

$$\left.\begin{aligned}\alpha_0&=pm-\omega\\ \beta_0&=qm+\omega\end{aligned}\right\}\quad \text{avec } |\omega|<1.$$

Si m tend vers l'infini, on a

$$\lim\frac{\alpha_0}{m}=p,\qquad \lim\frac{\beta_0}{m}=q\qquad (m\to\infty)$$

Je dis maintenant que, lorsque m augmente indéfiniment, la probabilité la plus grande $P(\alpha_0)$ tend vers zéro.

Pour faire cette démonstration, je m'appuie sur un résultat d'analyse mathématique, qui sera démontré dans une Note à la fin de cet opuscule. Ce résultat concerne l'évaluation de factorielle de n quand n est très grand, on a la formule de *Stirling*

$$n!=1.2.3\ldots n=n^n e^{-n}\sqrt{2\pi n}(1+\varepsilon),$$

ε tendant vers zéro quand n augmente indéfiniment.

Admettons cette formule et évaluons $P(\alpha_0)$.

On a

$$P(\alpha_0)=\frac{(\alpha_0+\beta_0)!}{\alpha_0!\,\beta_0!}p^{\alpha_0}q^{\beta_0},$$

d'où, en employant le signe $\sim$ pour indiquer une valeur asymptotique

$$P(\alpha_0)\sim\frac{(\alpha_0+\beta_0)^{\alpha_0+\beta_0}e^{-(\alpha_0+\beta_0)}\sqrt{2\pi(\alpha_0+\beta_0)}\,p^{\alpha_0}q^{\beta_0}}{\alpha_0^{\alpha_0}e^{-\alpha_0}\sqrt{2\pi\alpha_0}\,\beta_0^{\beta_0}e^{-\beta_0}\sqrt{2\pi\beta_0}}.$$

Or on a

$$\left(\frac{\alpha_0+\beta_0}{\alpha_0}\right)^{\alpha_0} p^{\alpha_0} = \left(\frac{m}{\alpha_0}\right)^{\alpha_0} p^{\alpha_0} \sim 1,$$

$$\left(\frac{\alpha_0+\beta_0}{\beta_0}\right)^{\beta_0} q^{\beta_0} = \left(\frac{m}{\beta_0}\right)^{\beta_0} q^{\beta_0} \sim 1,$$

$$\sqrt{\frac{1}{\alpha_0}+\frac{1}{\beta_0}} = \sqrt{\frac{m}{\alpha_0\beta_0}} = \frac{1}{\sqrt{m}}\,\frac{1}{\sqrt{\frac{\alpha_0}{m}\,\frac{\beta_0}{m}}} \sim \frac{1}{\sqrt{pqm}},$$

d'où

$$P(\alpha_0) = \frac{1}{\sqrt{2\pi pqm}}(1+\eta) \qquad (\eta \to 0 \text{ pour } m \to \infty).$$

On voit que $P(\alpha_0)$ tend vers zéro comme $\frac{1}{\sqrt{m}}$, lorsque m croît indéfiniment.

Je considère enfin, au lieu du terme de rang α_0, le terme de rang $\alpha_0 - i$ et je me propose encore de l'évaluer pour de très grandes valeurs de m. Alors α_0 sera aussi un grand nombre de l'ordre de m. Je ne supposerai pas que i demeure un nombre fini, je me réserve le droit de considérer des valeurs de i très grandes, mais telles cependant que $\frac{i}{\sqrt{\alpha_0}}$ reste fini. Il en sera de même de $\frac{i}{\sqrt{\beta_0}}$. Comme conséquence, les termes en $\frac{i^3}{\alpha_0^2} = \left(\frac{i^2}{\alpha_0}\right)^2 \frac{1}{i}$ et $\frac{i^3}{\beta_0^2} = \left(\frac{i^2}{\beta_0}\right)^2 \frac{1}{i}$ tendront vers zéro et pourront être négligés,

Je dis qu'on a alors

$$P(\alpha) = P(\alpha_0 - i) \sim \frac{1}{\sqrt{2\pi pqm}} e^{-\frac{i^2}{2pqm}}.$$

Le calcul est tout semblable au précédent. L'on a

$$P(\alpha) = P(\alpha_0 - i) = \frac{(\alpha_0+\beta_0)!}{(\alpha_0-i)!\,(\beta_0+i)!} p^{\alpha_0-i} q^{\beta_0+i},$$

d'où, pour la valeur asymptotique,

$$P(\alpha_0 - i) \sim \frac{(\alpha_0+\beta_0)^{\alpha_0+\beta_0} e^{-(\alpha_0+\beta_0)} \sqrt{2\pi(\alpha_0+\beta_0)}\, p^{\alpha_0-i} q^{\beta_0+i}}{(\alpha_0-i)^{\alpha_0-i} e^{-(\alpha_0-i)} \sqrt{2\pi(\alpha_0-i)}\,(\beta_0+i)^{\beta_0+i} e^{-(\beta_0+i)} \sqrt{2\pi(\beta_0+i)}}$$

ou

$$P(\alpha_0 - i) \sim \frac{1}{\sqrt{2\pi}} \sqrt{\frac{1}{\alpha_0 - i}+\frac{1}{\beta_0+i}}\, \frac{(\alpha_0+\beta_0)^{\alpha_0+\beta_0}}{(\alpha_0-i)^{\alpha_0-i}(\beta_0+i)^{\beta_0+i}} p^{\alpha_0-i} q^{\beta_0+i}.$$

Comme je suppose que $\frac{i}{\sqrt{\alpha_0}}$ et $\frac{i}{\sqrt{\beta_0}}$ restent finis et que α_0 et β_0 augmentent indéfiniment, les fractions $\frac{i}{\alpha_0}$ et $\frac{i}{\beta_0}$ tendent vers zéro, et l'on a

$$\sqrt{\frac{1}{\alpha_0 - i} + \frac{1}{\beta_0 + i}} = \sqrt{\frac{\alpha_0 + \beta_0}{(\alpha_0 - i)(\beta_0 + i)}} \sim \frac{1}{\sqrt{mpq}}.$$

D'autre part, nous avons vu qu'on peut poser

$$p = \frac{\alpha_0 - \omega}{m}, \qquad q = \frac{\beta_0 + \omega}{m} \qquad (|\omega| < 1)$$

ou

$$p = \frac{\alpha_0}{\alpha_0 + \beta_0}\left(1 - \frac{\omega}{\alpha_0}\right),$$

$$q = \frac{\beta_0}{\alpha_0 + \beta_0}\left(1 + \frac{\omega}{\beta_0}\right),$$

d'où

$$\frac{(\alpha_0 + \beta_0)^{\alpha_0 + \beta_0}}{(\alpha_0 - i)^{\alpha_0 - i}(\beta_0 + i)^{\beta_0 + i}} p^{\alpha_0 - i} q^{\beta_0 + i}$$
$$= \frac{1}{\left(1 - \frac{i}{\alpha_0}\right)^{\alpha_0 - i}\left(1 + \frac{i}{\beta_0}\right)^{\beta_0 + i}} \left(1 - \frac{\omega}{\alpha_0}\right)^{\alpha_0\left(1 - \frac{i}{\alpha_0}\right)} \left(1 + \frac{\omega}{\beta_0}\right)^{\beta_0\left(1 + \frac{i}{\beta_0}\right)}.$$

Mais si ω tend vers une limite ω_0, on a

$$\lim\left[\left(1 - \frac{\omega}{\alpha_0}\right)^{\alpha_0}\right]^{1 - \frac{i}{\alpha_0}} = e^{-\omega_0},$$

$$\lim\left[\left(1 + \frac{\omega}{\beta_0}\right)^{\beta_0}\right]^{1 + \frac{i}{\beta_0}} = e^{+\omega_0},$$

et le produit des deux limites est égal à l'unité.

Il reste à calculer le produit

$$\left(1 - \frac{i}{\alpha_0}\right)^{\alpha_0 - i}\left(1 + \frac{i}{\beta_0}\right)^{\beta_0 + i},$$

c'est ce qu'il est facile de faire en prenant le logarithme et en se

servant du développement connu de $\log(1+x)$. J'ai, en effet,

$$\log\left(1-\frac{i}{\alpha_0}\right)^{\alpha_0-i} = (\alpha_0-i)\left(-\frac{i}{\alpha_0}-\frac{i^2}{2\alpha_0^2}-\frac{i^3}{3\alpha_0^3}-\ldots\right)$$

$$= -i+\frac{i^2}{2\alpha_0}+\frac{i^3}{6\alpha_0^2}+\ldots,$$

$$\log\left(1+\frac{i}{\beta_0}\right)^{\beta_0+i} = (\beta_0+i)\left(\frac{i}{\beta_0}-\frac{i^2}{2\beta_0^2}+\frac{i^3}{3\beta_0^3}-\ldots\right)$$

$$= i+\frac{i^2}{2\beta_0}-\frac{i^3}{6\beta_0^2}+\ldots$$

En faisant la somme et négligeant les termes en $\frac{i^3}{\alpha_0^2}$, $\frac{i^3}{\beta_0^2}$ et suivants, le second membre devient

$$\frac{i^2}{2}\left(\frac{1}{\alpha_0}+\frac{1}{\beta_0}\right) \sim \frac{i^2}{2mpq}$$

et par suite

$$\left(1-\frac{i}{\alpha_0}\right)^{\alpha_0-i}\left(1+\frac{i}{\beta_0}\right)^{\beta_0+i} \sim \frac{i^2}{2mpq}.$$

On en déduit aussitôt

$$P(\alpha_0-i) \sim \frac{1}{\sqrt{2\pi pqm}} e^{-\frac{i^2}{2mpq}}.$$

Il résulte immédiatement de cette formule que les valeurs asymptotiques de $P(\alpha_0+i)$ et de $P(\alpha_0-i)$ sont les mêmes.

Répétition des épreuves. — Le jeu que nous avions imaginé consistait soit à faire tourner un disque partagé en deux secteurs d'amplitudes p et q, soit à tirer une boule d'une urne qui contient une proportion p de boules blanches et une proportion q de boules noires, puis à répéter l'épreuve un nombre m de fois dans les mêmes conditions, par exemple en remettant à chaque fois dans l'urne la boule qui vient d'être tirée; mais de plus nous avions figuré l'enregistrement du résultat de ces épreuves répétées, en marquant les coups par le déplacement d'une pièce qui se déplace sur un damier d'un pas suivant Ox dans l'une des éventualités et d'un pas suivant Oy dans l'autre éventualité. Si la pièce, partant de la case $(1,1)$ arrive sur la case $(\alpha+1,\ \beta+1)$, c'est que sur les $(m=\alpha+\beta)$ coups, il y en a eu α favorables à l'événement E dont la probabilité est p et β favorables à l'événement contraire.

En ce cas, la *fréquence* f de l'événement E a été, par définition, $\frac{\alpha}{\alpha+\beta}\left(\text{ou } \frac{\alpha}{m}\right)$.

On peut se demander si, à la longue, la fréquence f ne va pas tendre vers p.

Évidemment, il ne peut s'agir ici que d'une question de probabilité, car il est possible, quoique peu vraisemblable, que jamais il n'y ait un coup favorable à E.

D'ailleurs, la probabilité de la fréquence $\frac{\alpha}{m}$ devient maxima, quand m croît indéfiniment, pour une valeur α_0 telle que $\frac{\alpha_0}{m}$ tende vers p, mais cette probabilité maxima tend elle-même vers zéro.

Si l étant un entier fini, j'envisageais toutes les fréquences de $\frac{\alpha_0 - l}{m}$ à $\frac{\alpha_0 + l}{m}$ la probabilité serait

$$P = P(\alpha_0 - l) + P(\alpha_0 - l + 1) + \ldots + P(\alpha_0 + l - 1) + P(\alpha_0 + l).$$

Mais chacune de ces probabilités tendant vers zéro quand m augmente indéfiniment la probabilité P serait encore infiniment petite.

Pour avoir une probabilité finie, il faut donc que le champ ($\alpha_0 - l$ à $\alpha_0 + l$) augmente indéfiniment. La fréquence serait alors comprise entre $\frac{\alpha_0 - l}{m}$ et $\frac{\alpha_0 + l}{m}$ et comme $\frac{\alpha_0}{m}$ tend vers p, la fréquence tendra aussi vers p, si l'on fait croître l assez lentement avec m, pour que $\frac{l}{m}$ tende tout de même vers zéro.

C'est en particulier une condition qui sera satisfaite si $\frac{l}{\sqrt{\alpha_0}}$ ou $\frac{l}{\sqrt{m}}$ reste fini, quoique l augmente indéfiniment avec m.

On peut d'ailleurs imaginer que $\frac{l}{\sqrt{m}}$ soit susceptible de prendre des valeurs aussi grandes que l'on voudra et que cependant $\frac{l}{m}$ tende vers zéro. En posant, par exemple, $l = m^{\frac{2}{3}}$, on aurait :

$$\frac{l}{\sqrt{m}} = m^{\frac{1}{6}} \to \infty \qquad \text{et} \qquad \frac{l}{m} = m^{-\frac{1}{3}} \to 0.$$

Envisageons pour champ attribué au nombre variable i l'intervalle qui va de l_1 à l_2, alors la probabilité P sera donnée par la

somme

$$P = P(\alpha_0 + l_1) + P(\alpha_0 + l_1 + 1) + \ldots + P(\alpha_0 + l_2).$$

Si nous supposons que m et α_0 augmentent indéfiniment et que pour toute valeur de i considérée dans l'intervalle en question $\frac{i}{\sqrt{\alpha_0}}$ conserve une valeur finie, nous aurons, pour le terme général de cette somme

$$P(\alpha_0 + i) \sim \frac{1}{\sqrt{2\pi pqm}} e^{-\frac{i^2}{2pqm}}.$$

Cela étant, je pose

$$x = \frac{i}{\sqrt{pqm}} = \frac{\alpha - \alpha_0}{\sqrt{pqm}}, \qquad x_1 = \frac{l_1}{\sqrt{pqm}}, \qquad x_2 = \frac{l_2}{\sqrt{pqm}}.$$

On appelle *écart type* l'expression $\sqrt{pqm}$; et si je désigne par Δx l'accroissement de x quand α ou i augmentent d'une unité, j'aurai

$$\Delta x = \frac{1}{\sqrt{pqm}}.$$

Alors il viendra

$$P \sim \sum_{x=x_1}^{x=x_2} \Delta x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Mais comme m est infiniment grand, Δx est infiniment petit et nous pouvons écrire

$$P \sim \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx.$$

Telle est la valeur asymptotique de P, lorsque x_1 et x_2 sont finis.

Cette intégrale a alors une valeur finie et déterminée. Ce sera la probabilité pour que, le nombre m des épreuves augmentant indéfiniment, la fréquence de l'arrivée effective de l'événement E soit comprise entre $\frac{\alpha_0 - l_1}{m}$ et $\frac{\alpha_0 - l_2}{m}$.

Nous avons vu que l'on a

$$\alpha_0 = mp - \omega, \qquad \beta_0 = mq + \omega.$$

Donc, comme ω est, en valeur absolue, inférieure à l'unité, on pourra remplacer sans erreur α_0 et β_0 par leurs valeurs approchées mp et mq, l'erreur relative étant infiniment petite.

En particulier, nous avons posé

$$x = \frac{\alpha - \alpha_0}{\sqrt{pqm}} = \frac{\beta_0 - \beta}{\sqrt{pqm}}.$$

On peut, au contraire, au lieu de la variable x, introduire la variable λ qui est donnée par les relations

$$\lambda = \frac{\alpha - mp}{\sqrt{2pqm}} = \frac{mq - \beta}{\sqrt{2pqm}}.$$

Cette grandeur λ porte le nom d'*écart réduit* et quand m est très grand, on a sensiblement

$$\lambda = \frac{x}{\sqrt{2}}.$$

Soit donc

$$\alpha = mp + \lambda\sqrt{2pqm},$$
$$\beta = mq - \lambda\sqrt{2pqm}.$$

Supposons λ positif, alors on a (en se rappelant que $p + q$ est égal à 1)

$$\frac{\alpha}{\beta} - \frac{p}{q} = \frac{\alpha q - \beta p}{\beta q} = \frac{\lambda\sqrt{2pqm}}{\beta q} > 0$$

et cette différence tendra vers zéro, pour une valeur de λ donnée, puisque $\frac{\beta}{\sqrt{m}}$ tend vers l'infini avec m.

Changeons maintenant λ en $-\lambda$, et écrivons

$$\alpha = mp - \lambda\sqrt{2pqm},$$
$$\beta = mq + \lambda\sqrt{2pqm}.$$

C'est maintenant la différence

$$\frac{p}{q} - \frac{\alpha}{\beta} = \frac{\beta p - \alpha q}{q\beta} = \frac{\lambda\sqrt{2pqm}}{\beta q},$$

qui est positive, et elle tend vers zéro en même temps que $\frac{\sqrt{m}}{\beta}$, pour m très grand.

La condition pour que la valeur absolue de

$$\frac{\alpha}{\beta} - \frac{p}{q}$$

soit supérieure à un nombre positif donné ε, si petit qu'il soit, s'écrit donc

$$\lambda\sqrt{2pqm} > \varepsilon q(mq \mp \lambda\sqrt{2pqm})$$

ou

$$\lambda > \frac{\varepsilon q^2\sqrt{m}}{(1 \pm \varepsilon q)\sqrt{2pq}}$$

et l'on voit que λ devra croître indéfiniment, si m est infiniment grand, quelle que soit la valeur très petite, mais fixe, que l'on ait assignée à ε.

Nous voyons donc, d'abord que, lorsque l'on se donne l'écart réduit, le rapport des fréquences $\frac{\alpha}{m}$ et $\frac{\beta}{m}$ tend vers le rapport des probabilités p, q, quand le nombre des épreuves répétées augmente indéfiniment.

D'autre part, quand le nombre des épreuves augmente indéfiniment, l'écart réduit doit augmenter indéfiniment de part et d'autre de son zéro pour que la différence entre le rapport des fréquences et celui des probabilités élémentaires reste supérieure à un nombre donné ε, aussi petit qu'on veut.

D'une façon précise, si je pose

$$\lambda' = \frac{\varepsilon q^2\sqrt{m}}{(1+\varepsilon q)\sqrt{2pq}}, \qquad \lambda'' = \frac{\varepsilon q^2\sqrt{m}}{(1-\varepsilon q)\sqrt{2pq}},$$

il faudra qu'on ait

$$\lambda > \lambda' \quad \text{ou} \quad -\lambda < -\lambda''.$$

La probabilité pour que l'on ait

$$\left|\frac{\alpha}{\beta} - \frac{p}{q}\right| > \varepsilon,$$

est donc

$$1 - \frac{1}{\sqrt{\pi}}\int_{-\lambda''}^{\lambda'} e^{-\lambda^2}\,d\lambda.$$

Or quand λ' et λ'' augmentent indéfiniment, l'intégrale tend vers $\sqrt{\pi}$ et la probabilité en question tend vers zéro.

Donc la probabilité pour que la différence entre le rapport des fréquences et celui des probabilités élémentaires, c'est-à-dire entre le rapport des fréquences et le rapport théorique $\frac{p}{q}$, soit

supérieur à un nombre donné ε tend vers zéro quand le nombre des épreuves augmente indéfiniment.

On peut dire encore :

Si p est la probabilité de l'événement E et q celle de son contraire dans une épreuve déterminée, si α est le nombre de fois que l'événement E se présente quand on répète m fois ladite épreuve, si j'appelle P la probabilité pour que la différence entre p et $\frac{\alpha}{m}$ soit en valeur absolue inférieure à un nombre donné, on peut toujours prendre m assez grand pour que P diffère de l'unité aussi peu que l'on voudra.

C'est cette proposition qu'on désigne sous le nom de Théorème de Bernoulli.

En résumé, si le champ de variation que nous adoptons pour l'écart i est compris dans l'intervalle fini et déterminé $(\alpha_0 - l, \alpha_0 + l)$, la probabilité pour que l'écart effectif i soit réellement compris dans cet intervalle tend vers zéro quand le nombre des épreuves réitérées augmente indéfiniment.

Si le champ de variation considéré augmente indéfiniment, il convient au lieu de considérer l'écart i de recourir à la variable x ou à l'écart réduit λ. Si λ, qui est égal à $\frac{\alpha - mp}{\sqrt{2pqm}}$ reste compris entre les deux limites λ_1 et $-\lambda_1$ finies et déterminées, cela implique que l'écart $i = \alpha - \alpha_0$ croît indéfiniment et que son ordre de grandeur est celui de $\sqrt{m}$; alors la probabilité pour que λ soit compris dans l'intervalle considéré a une valeur finie dont l'expression analytique est

$$P = \frac{1}{\sqrt{\pi}} \int_{-\lambda_1}^{+\lambda_2} e^{-\lambda} \, d\lambda.$$

Cette probabilité augmente constamment avec λ_1 et tend vers l'unité quand λ_1 augmente indéfiniment.

Si, d'autre part, nous considérons l'écart maximum ε entre $\frac{\alpha}{\beta}$ et $\frac{p}{q}$, pour les valeurs croissantes de m, nous voyons que si λ_1 est donné, ε décroît lorsque m augmente indéfiniment, d'autre part si ε est donné il faut que λ_1 soit supérieur à une certaine limite et quand ε est pris de plus en plus petit, λ_1 doit croître indéfiniment.

pour un ε donné, et cela comme $\sqrt{m}$. Par suite, λ_1 croissant indéfiniment, P tend vers 1.

8. **Lois de probabilités continues.** — Nous avons considéré un jeu dans lequel, à chaque coup, la probabilité de l'événement E est p, tandis que celle de l'événement contraire est q, de sorte que l'on a

$$p + q = 1$$

et nous avons imaginé que le nombre des épreuves m augmentait indéfiniment.

C'est dans ce passage à l'infini que nous avons été amené à envisager l'écart réduit λ et à lui attribuer une loi de probabilité continue, en ce sens que la probabilité pour que l'écart réduit soit compris entre λ et $\lambda + d\lambda$ est égale à

$$\frac{1}{\sqrt{\pi}} e^{-\lambda^2} d\lambda.$$

Si je porte λ en abscisse et

$$y = \frac{1}{\sqrt{\pi}} e^{-\lambda^2}$$

en ordonnées, j'obtiens une courbe qu'on appelle ordinairement la courbe en cloche et cette fonction y de λ caractérise la densité de la probabilité dans la loi de probabilité à laquelle on donne le nom de *loi de Gauss*. Le problème de probabilités discontinues a été remplacé par un problème du continu. Malgré les approximations faites, on remarquera que la somme des probabilités élementaires est rigoureusement égale à 1.

Considérons donc la loi de Gauss. Je dis que la probabilité d'un écart relatif λ supérieur à un nombre a est infiniment petite quand a est infiniment grand.

Cette probabilité P est, en effet, égale à

$$P = \frac{1}{\sqrt{\pi}} \int_a^{+\infty} e^{-\lambda^2} d\lambda.$$

Je vais, pour *calculer* P, donner le développement asymptotique de cette intégrale. Dans ce but, par le moyen du changement de variable

$$\lambda = a + \frac{u}{2a},$$

je ramène l'intervalle d'intégration à être l'intervalle o, $+\infty$ et j'ai

$$\sqrt{\pi}\,P = \frac{e^{-a^2}}{2a}\int_0^{+\infty} e^{-\frac{u^2}{4a^2}} e^{-u}\,du$$

et je développe $e^{-\frac{u^2}{4a^2}}$ par la formule de Mac Laurin. J'ai

$$e^{-\frac{u^2}{4a^2}} = \sum_{\nu=0}^{n-1} (-1)^\nu \frac{u^{2\nu}}{(2a)^{2\nu}\nu!} + (-1)^n \left(\frac{u^2}{4a^2}\right)^n \frac{1}{n!} e^{-\theta\frac{u^2}{4a^2}} \quad (0 < \theta < 1).$$

D'autre part, j'ai

$$\Gamma(2\nu+1) = (2\nu)! = \int_0^\infty u^{2\nu} e^{-u}\,du.$$

Tenant compte de ces deux résultats, les intégrations peuvent s'effectuer et j'obtiens

$$\sqrt{\pi}\,P = e^{-a^2}\sum_{\nu=0}^{n-1}(-1)^\nu \frac{(2\nu)!}{\nu!\,(2a)^{2\nu+1}} + R_n$$

avec l'inégalité

$$R_n < \frac{(2n)!}{n!}\,\frac{e^{-a^2}}{(2a)^{2n+1}}.$$

Si donc, je calcule un nombre de termes fini, n étant ainsi déterminé, R_n et les termes intégrés tendront vers zéro quand a augmentera indéfiniment.

La série limitée aux termes intégrés donnera donc une approximation de $P\sqrt{\pi}$ et le reste R_n ne dépassera pas une certaine valeur qui pourra être estimée et qui sera petite si a est suffisamment grand.

Mais R_n est une fonction, non seulement de a, mais encore de n et si, au lieu de se donner n, on se fixe une valeur de a, le reste R_n ne tendra pas vers zéro, car la série formée par les termes intégrés est divergente, comme on s'en rend compte aisément en prenant le rapport d'un terme au précédent.

On aurait pour P une approximation moins satisfaisante en écrivant simplement

$$\sqrt{\pi}\,P = \int_a^\infty e^{-\lambda^2}\,d\lambda < \int_a^\infty e^{-\lambda}\,d\lambda = e^{-a}.$$

Cette inégalité suffirait pourtant à montrer que P tend vers zéro

quand a augmente indéfiniment. Mais la série se prête au calcul exact de P.

9. **Loi de Poisson.** — Nous savons que si l'on répète m fois une épreuve où l'événement E a la probabilité p, et son contraire la probabilité q, la probabilité $P(\alpha)$ pour que, sur les m épreuves, l'événement E se soit produit réellement un nombre de fois égal à α est donnée par la formule

$$P(\alpha) = \frac{m!}{\alpha!(m-\alpha)!} p^{\alpha}(1-p)^{m-\alpha}.$$

On arrive à la loi de probabilité dite de Poisson en supposant que la probabilité p devient infiniment petite en même temps que le nombre des épreuves devient infiniment grand, de telle façon que l'on ait

$$\lim mp = y,$$

y ayant une valeur finie et déterminée.

C'est encore la formule de Stirling pour l'approximation des factorielles qu'il y a lieu d'appliquer.

On a, tout d'abord, sensiblement en remplaçant p par $\frac{y}{m}$

$$p^{\alpha}(1-p)^{m-\alpha} = \frac{y^{\alpha}(m-y)^{m-\alpha}}{m^m}.$$

Puis, dans les mêmes conditions

$$P(\alpha) = \frac{m^m e^{-m}\sqrt{2\pi m}\, y^{\alpha}(m-y)^{m-\alpha}}{\alpha!\,(m-\alpha)^{m-\alpha} e^{-m+\alpha}\sqrt{2\pi(m-\alpha)}\, m^m}$$

ou

$$P(\alpha) = \sqrt{\frac{m}{m-\alpha}}\left(\frac{m-y}{m-\alpha}\right)^{m-\alpha} \frac{y^{\alpha}}{\alpha!\, e^{-\alpha}}.$$

Mais on a

$$\lim_{m \to \infty} \sqrt{\frac{m}{m-\alpha}} = 1,$$

$$\lim_{m \to \infty} \left(\frac{m-y}{m-\alpha}\right)^{m-\alpha} = \lim_{m \to \infty} \left(1 + \frac{\alpha - y}{m-\alpha}\right)^{m-\alpha} = e^{\alpha-y},$$

d'où

$$P(\alpha) = \frac{y^{\alpha} e^{-y}}{\alpha!}.$$

Cette loi de probabilité est appelée *loi de Poisson*.

On sait que si α_0 est le nombre de fois le plus probable, on a

$$\alpha_0 = mp - \omega \qquad (|\omega| < 1);$$

or on a

$$y = \lim mp.$$

D'où

$$\lim \frac{y}{\alpha_0} = 1 \qquad \text{pour } m \to \infty.$$

Telle est la signification de y.

Scintillations. — Quand on place en regard d'un écran phosphorescent un grain de substance radioactive, on voit des scintillations sur l'écran. Elles sont produites par le choc des particules positives produites pendant la désintégration de la substance radioactive, phénomène qui semble s'effectuer au hasard.

Prenons comme unité de temps l'intervalle où il se produit en moyenne une scintillation, $P(\alpha)$ sera la probabilité pour qu'il s'en produise α dans l'intervalle de temps y.

Cette interprétation résultera plus clairement de la présentation que nous allons faire maintenant de la loi de Poisson, en suivant M. Borel.

REPRÉSENTATION GÉOMÉTRIQUE DE LA LOI DE POISSON.

Sachant que sur une droite indéfinie dans les deux sens $X'X$, il y a une infinité de points $A_1, A_2, \ldots, A_\alpha, \ldots$ répartis au hasard, mais cependant de telle façon qu'il y en ait en moyenne un par unité de longueur, on demande la probabilité $P(\alpha)$ qu'il y en ait α sur un segment de longueur y (*Cf.* p. 60 — 2°).

Pour fixer les idées, considérons d'abord un intervalle fini l sur lequel m points sont répartis au hasard, il y en aura en moyenne un nombre $\frac{m}{l}$ par unité de longueur. Dans notre cas, nous ferons tendre m et l vers l'infini de telle façon qu'on ait

$$\lim \frac{m}{l} = 1.$$

Et nous aurons alors un nombre infini de points répartis de manière qu'il y en ait un *en moyenne* par unité de longueur. La longueur y deviendra ainsi un pur nombre.

Considérons encore l'intervalle fini l et, dans cet intervalle, le segment y. Et supposons qu'on ait à poser au hasard sur l'intervalle l un point Q, la probabilité pour que ce point soit placé sur le segment y sera $\frac{y}{l}$. C'est ainsi que nous faisons jouer le hasard. On voit que si l tend vers l'infini en grandeur, tandis que y reste constant, cette probabilité tend vers zéro.

Continuons à envisager l'intervalle fini l et un nombre m de points matériels, individualisés par leurs numéros et que nous appellerons $A_1, A_2, \ldots, A_m$. A l'intérieur de l'intervalle l se trouve le segment y et nous répartissons au hasard sur l successivement les points A_1, A_2 etc. Soit ϖ_α la probabilité pour que les points $A_1, A_2, \ldots, A_\alpha$ soient sur le segment y, tandis que les points $A_{\alpha+1}, A_{\alpha+2}, \ldots, A_m$ seraient sur le segment complémentaire $l-y$. Quelle différence y-a-t-il entre ce cas et celui où le point A_α, au lieu d'être sur le segment y, était sur le segment complémentaire ? C'est que la probabilité a joué en faveur du point A_α. Si donc j'appelle $\varpi_{\alpha-1}$ la probabilité pour que les points A_1, A_2, $\ldots$, $A_{\alpha-1}$ soient sur le segment y tandis que A_α et les points suivants d'indice supérieur sont sur le segment complémentaire, j'ai

$$\frac{\varpi_\alpha}{y} = \frac{\varpi_{\alpha-1}}{l-y}.$$

D'autre part, si nous effaçons les numéros qui distinguent les points matériels et si nous n'avons égard qu'au nombre de points qui se trouvent sur le segment y, sans avoir à distinguer l'individualité de chacun, nous aurons pour la probabilité $p(\alpha)$ qu'il y ait α points sur y, la valeur suivante obtenue en réunissant les cas d'égale probabilité :

$$p(\alpha) = C_m^\alpha \varpi(\alpha),$$

d'où

$$\frac{p(\alpha)}{p(\alpha-1)} = \frac{\dfrac{m!}{\alpha!(m-\alpha)!}}{\dfrac{m!}{(\alpha-1)!(m-\alpha+1)!}} \frac{\varpi(\alpha)}{\varpi(\alpha-1)}$$

$$= \frac{m-\alpha+1}{\alpha} \frac{y}{l-y} = \frac{ym}{\alpha l} \frac{1-\dfrac{\alpha-1}{m}}{1-\dfrac{y}{l}}.$$

Supposons maintenant, pour revenir aux conditions de l'énoncé,

que m et l tendent vers l'infini, de façon que leur rapport tende vers l'unité et que y reste constant, ainsi que le nombre α qu'on s'est donné, alors $p(\alpha)$ tendra vers la probabilité cherchée $P(\alpha)$ et l'on aura la loi de récurrence

$$\frac{P(\alpha)}{P(\alpha-1)}=\frac{y}{\alpha}, \quad \frac{P(1)}{P(0)}=y.$$

D'où

$$P(\alpha)=\frac{y^\alpha}{\alpha!}P(0).$$

Pour déterminer $P(0)$, appliquons le principe de la probabilité totale, il vient

$$1=\sum_{\alpha=0}^{\infty}P(\alpha)=P_0\sum\frac{y^\alpha}{\alpha!}=P_0\,e^y,$$

d'où

$$P_0=e^{-y}$$

et ceci donne un sens intéressant à y. Enfin, on a

$$P(\alpha)=\frac{y^\alpha e^{-y}}{\alpha!},$$

ce qui donne la loi de Poisson.

10. **Probabilités géométriques.** — On prend un point au hasard sur l'axe des x, à l'intérieur d'un segment de longueur égale à l'unité. La probabilité que le point soit dans l'intervalle dx est

$$f(x)\,dx.$$

Effectuons un changement de variable, la probabilité deviendra

$$g(x')\,dx'.$$

Si l'on veut que la probabilité reste inchangée quand on fait subir à tous les points une translation, il faut que le changement de variable

$$x'=x+a$$

ne modifie pas la probabilité. Or, on a

$$dx'=dx.$$

Donc il faut qu'on ait $f(x)=g(x')$ et cela pour une infinité de valeurs de x, donc $f(x)$ et $g(x')$ se réduisent à une constante, qui devra être égale à l'unité.

Semblablement, prenons un point au hasard dans un plan sur une surface égale à l'unité. La probabilité élémentaire pourra être prise égale à

$$dx\,dy$$

en coordonnées rectangulaires, car elle sera ainsi indépendante d'un déplacement arbitraire consistant en une translation suivie d'une rotation. Si l'on pose

$$x' = x + a, \qquad y' = y + b$$

ou bien

$$x' = x\cos\alpha - y\sin\alpha,$$
$$y' = x\sin\alpha + y\cos\alpha.$$

Dans les deux cas le déterminant fonctionnel est égal à l'unité et l'on a

$$dx'\,dy' = dx\,dy,$$

ce qui justifie l'assertion.

Si l'on pose

$$x = \rho\cos\varphi, \qquad y = \rho\sin\varphi,$$

on a

$$dx = d\rho\cos\varphi - \rho\sin\varphi\,d\varphi,$$
$$dy = d\rho\sin\varphi + \rho\cos\varphi\,d\varphi,$$

d'où un déterminant fonctionnel égal à ρ et l'on a

$$dx\,dy = \rho\,d\rho\,d\varphi.$$

Par suite, en coordonnées polaires, on pourra prendre $\rho\,d\rho\,d\varphi$ pour probabilité élémentaire, c'est-à-dire pour la probabilité que le point choisi soit au voisinage du point ρ, φ à l'intérieur d'un élément d'aire de surface $\rho\,d\rho\,d\varphi$. La grandeur de la surface élémentaire peut être prise pour évaluation de la probabilité, si celle-ci doit être indépendante d'un changement de coordonnées polaires.

Considérons maintenant une droite D, sous forme canonique,

$$x\cos\alpha + y\sin\alpha - p = 0.$$

Elle est déterminée par les deux paramètres α et p. Déplaçons cette droite. Cela peut se faire par une rotation autour de l'origine d'un angle φ; et l'on a

$$\alpha' = \alpha + \varphi,$$

cette rotation sera suivie d'une translation u, que je puis décom-

poser en deux, l'une suivant la perpendiculaire à la nouvelle position de la droite D et l'autre suivant la droite D elle-même. Cette dernière ne modifie pas la position de la droite; la première s'exprime par

$$p' = p + r,$$

r étant une constante qui avec φ détermine le déplacement de la droite. On a

$$d\alpha' = d\alpha, \qquad dp' = dp$$

et par suite

$$d\alpha'\, dp' = d\alpha\, dp.$$

Problème. — *On prend deux points au hasard sur la circonférence* (C), *quelle est la probabilité pour que la corde qui les joint soit plus grande que le côté du triangle équilatéral inscrit.*

Réponse. — Le premier point étant choisi arbitrairement, le second doit être pris sur le tiers de circonférence opposé à la position du premier point. La probabilité est donc $\frac{1}{3}$.

Autre problème. — *Je trace une droite* D *dont la distance* p *au centre est inférieure au rayon; cette droite étant ainsi tracée au hasard, quelle est la probabilité pour que la corde interceptée soit de longueur supérieure au côté du triangle équilatéral inscrit.*

Réponse. — La probabilité est $\frac{1}{2}$.

Problème. — *On prend au hasard deux points* P *et* Q *sur un segment* AB *de longueur* a. *Quelle est la probabilité que la distance entre les points* P *et* Q *soit inférieure à* b ?

Je pose

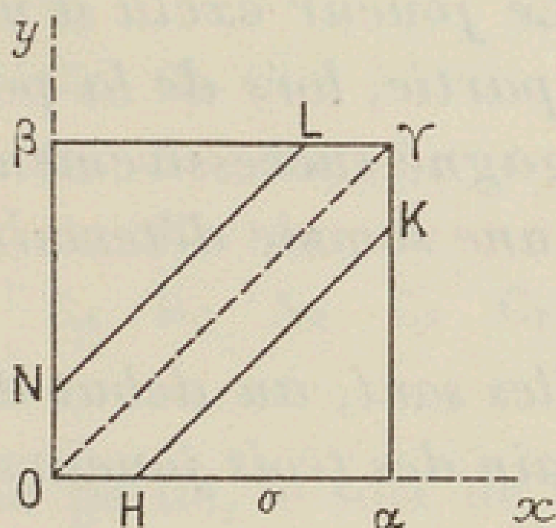
$$\mathrm{AP} = x, \qquad \mathrm{AQ} = y.$$

La probabilité pour que P tombe entre x et $x + dx$ est $\frac{dx}{a}$; celle pour que Q tombe entre y et $y + dy$ est $\frac{dy}{a}$. La probabilité composée de ces deux événements *indépendants* est

$$\frac{dx\, dy}{a^2}.$$

L'événement où AP est égal à x et où AQ est égal à y peut être représenté par un point M de coordonnées rectangulaires (x, y).

Fig. 1.



Ce point se trouve à l'intérieur du carré de côté a qui occupe l'angle interne des axes. Je trace les droites HK et NL qui ont pour équations

$$x - y = b, \qquad y - x = b.$$

Il est clair alors que la position favorable du point figuratif M est celle où il tombe à l'intérieur de la bande découpée dans le carré par ces deux droites et comme la probabilité élémentaire est l'élément de surface $dx\,dy$, la probabilité cherchée est le rapport

$$\frac{\text{aire OHK}\gamma\text{LM}}{\text{aire du carré}} = \frac{2b}{a} - \frac{b^2}{a^2} = 1 - \frac{(a-b)^2}{a^2}.$$

11. **Problème de la Poule.** — Commençons par faire une réflexion préliminaire d'ordre général.

Si l'événement attendu peut se produire dans un nombre de cas égal à n et si le nombre des éventualités à envisager, toutes de même probabilité, est égal à N, la probabilité de l'événement attendu est $\frac{n}{N}$ et cette probabilité sera nulle si n est égal à zéro; mais l'expression analytique de cette probabilité peut tendre vers zéro, sans que n soit nul, ce sera le cas, par exemple, si, n restant fini, N augmente indéfiniment.

Or, il y a des problèmes où le nombre de cas à envisager est infiniment grand.

C'est ce qui a lieu dans le problème de la Poule, qui s'énonce ainsi :

Énoncé. — *Trois joueurs* A, B, C *jouent à pile ou face, aux conditions suivantes :*

A *et* B *jouent une première partie,* C *remplace le perdant et la partie continue. Le joueur exclu d'une partie lutte contre le gagnant de cette partie, lors de la partie suivante, jusqu'à ce qu'un joueur ait gagné successivement les deux autres. Son gain est alors égal à une somme déterminée, qui est l'enjeu de la partie.*

On demande quelles sont, au début de la première partie, les probabilités de gain des trois joueurs.

On remarquera tout d'abord qu'il n'est pas certain que la partie se termine à l'avantage de l'un ou l'autre des joueurs, car il est évident que, quel que soit celui des deux joueurs A ou B qui gagne la première partie, il peut très bien se faire que, dans chacun des deux cas, aucun joueur ne réussisse à gagner deux fois de suite.

Cependant la somme des probabilités que chacun des trois joueurs a de gagner la partie est égale à l'unité, comme si l'on épuisait ainsi toutes les éventualités possibles, puisque la probabilité pour que le jeu continue indéfiniment sans vainqueur est exprimée par une fraction dont le numérateur est égal à 2 et dont le dénominateur augmente indéfiniment.

Si nous appelons x, y et z les chances respectives de A, B et C, l'application du principe de la probabilité totale nous donnera donc

$$x + y + z = 1$$

et d'autre part, il est évident que A et B ont des chances égales, d'où

$$x = y.$$

Il suffit donc de déterminer z, ou pour mieux dire $\frac{z}{2}$, en n'envisageant que l'hypothèse où A gagne la première partie. On aurait encore $\frac{z}{2}$, dans l'hypothèse où B la gagne.

Formons donc le tableau des éventualités possibles, en désignant par exemple par A_B une partie où les adversaires sont A et B et où A est vainqueur, une notation analogue étant employée dans tous les autres cas; on a alors la possibilité de ranger les diverses

éventualités dans un ordre bien déterminé, suivant que la partie se termine au second coup, ou au troisième, etc., ce qui correspond aux lignes successives du tableau ci-dessous

$$
\begin{array}{llllll}
A_B & A_C & & & & \\
A_B & C_A & C_B & & & \\
A_B & C_A & B_C & B_A & & \\
A_B & C_A & B_C & A_B & A_C & \\
A_B & C_A & B_C & A_B & C_A & C_B \\
\dots & \dots & \dots & \dots & \dots & \dots
\end{array}
$$

Au bout de la sixième partie, le sort de C est le même qu'au bout de la troisième. On voit donc que C peut gagner au bout d'un nombre de parties égal à 3, 6, 9, 12, etc. Et comme les probabilités de chaque coup de pile ou face sont toujours égales à $\frac{1}{2}$, on voit, en appliquant le théorème de la probabilité composée qu'on a

$$\frac{z}{2} = \frac{1}{2^3} + \frac{1}{2^6} + \frac{1}{2^9} + \dots,$$

d'où

$$= \frac{2}{7}$$

et par suite

$$x = y = \frac{5}{14}.$$

12. **Origine du calcul des probabilités.** — Le calcul des probabilités doit son origine à une controverse entre Pascal et Fermat au sujet de questions posées par le chevalier de Méré relativement à des jeux de hasard.

1° La première concerne la *règle des partis* ou de répartition. L'énoncé est le suivant :

Deux joueurs ont formé un fonds destiné à celui qui aura le plus tôt gagné trois parties; ils se séparent lorsque le premier en a gagné deux et le second une; on demande la part que chacun doit avoir sur le fonds du jeu, en supposant que la probabilité de gagner isolément une partie soit $\frac{1}{2}$ pour chaque joueur.

Solution. — S'il était joué une partie de plus et que le premier

joueur la gagnât, il aurait la mise totale; la probabilité de cet événement est $\frac{1}{2}$. Si cet événement n'arrivait pas, chacun des joueurs ayant gagné deux parties, la suivante déciderait nécessairement de leur sort; mais la probabilité qu'elle serait jouée étant $\frac{1}{2}$, on aura $\frac{1}{4}$ pour la probabilité qu'elle serait gagnée par le premier joueur, et de même pour le second. Or cette dernière probabilité est la seule que le second joueur ait de gagner, tandis que le premier avait $\frac{1}{2}$ dès la partie précédente; les probabilités *totales* sont par conséquent pour le premier joueur

$$\frac{1}{2}+\frac{1}{4}=\frac{3}{4}$$

et pour le second

$$\frac{1}{4}.$$

Si le fonds est de 64 pistoles, le premier en aura 48, le second 16.

2° La seconde question concerne le jeu des dés et s'énonce ainsi.

Trouver le nombre de jets de deux dés, dans lequel il y a autant de probabilité d'amener deux 6 (ou sonnez) que de ne pas le faire.

Solution. — La probabilité de *ne pas* amener le point sonnez est $q=\frac{35}{36}$ au premier coup; supposons donc qu'en effet le point sonnez n'ait pas été amené, alors il y aura encore la même probabilité q qu'il ne soit pas amené par le second jet, mais la probabilité que ce second jet ait lieu étant q, la probabilité de n'amener sonnez ni au premier jet, ni au second est q^2. On voit immédiatement que la probabilité de n'amener sonnez ni au premier jet, ni au second, etc., ni au $n^{\text{ième}}$ jet est q^n. Si donc le nombre n est tel qu'on ait

$$q^n>\frac{1}{2}, \qquad q^{n+1}<\frac{1}{2},$$

on voit qu'il y a au $n^{\text{ième}}$ jet plus de chance de ne pas avoir amené sonnez que de l'avoir amené et qu'au $(n+1)^{\text{ième}}$ jet, il y a au

contraire moins de chance. Résolvons donc l'équation en n ci-après

$$\left(\frac{35}{36}\right)^n = \frac{1}{2}, \qquad \text{d'où} \qquad n = 24,6.$$

On en conclut que l'événement proposé est moins probable que son contraire quand on n'embrasse que 24 jets et plus probable quand on en prend 25.

13. Utilité des méthodes de récurrence. — Considérons encore le problème des épreuves répétées.

Étant données m urnes identiques contenant chacune une proportion p de boules blanches et une proportion q de boules noires, sans autres boules, de sorte que $p+q$ est égal à 1, on tire une boule de chacune. Quelle est la probabilité P_α^m d'avoir amené α boules blanches exactement?

Appelons cet événement E_α^m. Or quand on a tiré une boule de chacune des $m-1$ premières urnes et que l'on procède au dernier tirage, l'événement peut se produire de deux façons; si l'événement au $m-1^{\text{ième}}$ tirage est E_α^{m-1}, il faut tirer une noire, si l'événement est $E_{\alpha-1}^{m-1}$, il faut tirer une blanche. Il n'y a pas d'autres éventualités à envisager, la première a la probabilité p, la seconde la probabilité q. En vertu de la théorie de la probabilité composée et de la probabilité totale, on a donc

$$P_\alpha^m = p P_{\alpha-1}^{m-1} + q P_\alpha^{m-1}.$$

Si donc on pose

$$P_\alpha^m = p^\alpha q^{m-\alpha} C_m^\alpha,$$

il vient, après division par le facteur commun $p^\alpha q^{m-\alpha}$,

$$C_m^\alpha = C_{m-1}^\alpha + C_{m-1}^{\alpha-1}.$$

C'est la relation de récurrence bien connue pour le nombre des combinaisons de m objets α à α.

Cette méthode paraît avoir été utilisée par les premiers auteurs qui ont traité le calcul des probabilités. A ce titre, elle mérite d'être rappelée (*Cf.* p. 20 — 3°).

14. Variable aléatoire. Coefficients caractéristiques. — Je me

propose ici de donner, pour fixer les idées, une représentation graphique d'une loi de probabilité.

Je commence par un exemple :

Soient p la probabilité d'un événement E, q la probabilité de l'événement contraire F; soit à calculer la probabilité pour que dans un nombre donné n d'épreuves l'événement E ait lieu α fois et F, $n-\alpha$ fois.

Nous avons trouvé, pour la probabilité, la valeur

$$P(\alpha) = \frac{\alpha!\,(n-\alpha)!}{n!} p^{\alpha} q^{n-\alpha}.$$

On peut réaliser l'expérience en imaginant un disque qui tourne devant un index et dont la surface est divisée en deux secteurs E et F, dont les surfaces ont pour mesures p et q, quand on prend la surface du disque pour unité. On fait tourner le disque n fois de suite et l'on gagne x_{α} francs, si sur les n épreuves, il y en a α dans lesquelles, à l'arrêt du disque, un secteur E se trouve devant l'index.

x_{α} serait négatif, s'il représentait une perte au lieu d'un gain. Nous supposons toutes ces valeurs x_{α} distinctes.

Alors je considère la somme gagnée x_{α} comme la valeur d'une variable aléatoire x et je la représente par un *segment* OX_{α} porté sur l'axe Ox, et j'affecte au point X_{α} une masse égale à $P(\alpha)$. Ceci étant fait pour les différentes valeurs de α, j'ai sur l'axe Ox une certaine répartition de masses, dont la somme est égale à l'unité, et cette répartition figure la loi de probabilité à laquelle est soumise la variable x. J'ai admis que les sommes gagnées x_{α} sont toutes distinctes. Dans le cas contraire, il faudrait supposer à l'abscisse correspondante une masse égale à la somme des probabilités amenant le gain x_{α}.

Au lieu de faire tourner le disque primitif n fois de suite, on pourrait imaginer un autre disque divisé en $n+1$ secteurs et qu'on ne ferait tourner qu'une fois, mais chacun des $n+1$ secteurs aurait une surface égale aux $n+1$ valeurs qu'est susceptible de prendre la probabilité $P(\alpha)$, quand α prend les valeurs $0, 1, \ldots, n$.

D'une façon générale, pour toute autre loi de probabilité, le disque est divisé en N secteurs de surfaces égales à $p_1, p_2, \ldots, p_N$,

ces nombres étant positifs et ayant pour somme l'unité; sous cette réserve, ils peuvent d'ailleurs être quelconques. On a donc

$$\sum_{i=1}^{N} p_i = 1.$$

Et à chaque point, d'abscisse x_i, est affectée une masse p_i. On appelle *espérance mathématique* le gain m_1 représenté par le centre de gravité des masses ainsi réparties. On a donc

$$m_1 \sum_{i=1}^{N} p_i = \sum_{i=1}^{N} p_i x_i$$

ou, plus simplement,

$$m_1 = \sum_{i=1}^{N} p_i x_i.$$

On dit aussi que m_1 est la *moyenne* du *premier ordre*. Comme d'après la loi des grands nombres, la fréquence réalisée se rapproche vraisemblablement de la probabilité théorique, on dit encore que m_1 est la *valeur probable*.

La *moyenne* du *second ordre* m_2 sera définie par l'équation analogue

$$m_2^2 = \sum_{i=1}^{N} p_i x_i^2,$$

m_2 correspondrait, en mécanique, au rayon de giration, la masse totale étant ici égale à l'unité.

La connaissance de m_2 a cet intérêt de nous faire savoir que les valeurs de x supérieures à m_2 sont peu probables.

Soient, en effet, $x'_1, x'_2, \ldots, x'_l$ les valeurs de x_i supérieures à km_2, le nombre k étant supposé grand. On aura, en appelant $p'_1, p'_2, \ldots, p'_l$ leurs probabilités $(l < \mathrm{N})$

$$m_2^2 > \sum_{j=1}^{l} p'_j x'^2_j > k^2 m_2^2 \sum_{j=1}^{l} p'_j.$$

D'où

$$\sum_{j=1}^{l} p'_j < \frac{1}{k^2}.$$

Or, le premier membre de cette inégalité n'est autre que la probabilité d'un gain supérieur à k fois la somme m_2 et par conséquent l'inégalité exprime que cette probabilité est infime, si k est un grand nombre.

On aurait de même pour la moyenne du troisième ordre m_3 la formule de définition

$$m_3^3 = \sum_i^N p_i x_i^3.$$

Revenons à la moyenne quadratique m_2. Portons l'origine au point d'abscisse ξ, ce qui veut dire que le gain x_i dont la probabilité était p_i est remplacé par une différence qui est le gain ou la perte $x_i - \xi$, avec la même probabilité.

Le nouveau gain quadratique moyen, est, dans ces nouvelles hypothèses, la quantité positive m'_2 définie par l'égalité

$$(m'_2)^2 = \sum_{i=1}^N p_i (x_i - \xi)^2,$$

Cherchons sa variation en fonction de ξ, on a

$$(m'_2)^2 = \sum_{i=1}^N p_i x_i^2 - 2\xi \sum_{i=1}^N p_i x_i + \xi^2 \sum p_i$$

ou

$$(m'_2)^2 = m_2^2 - 2 m_1 \xi + \xi^2 = m_2^2 - m_1^2 + (m_1 - \xi)^2.$$

D'où

$$(m'_2)^2 \geqq m_2^2 - m_1^2.$$

Le minimum de m'_2 étant obtenu pour $\xi = m_1$, je désigne ce minimum par μ_2, il a lieu quand on porte au centre de gravité l'origine des coordonnées, auquel cas le gain probable devient nul. On a donc

$$\mu_1 = \sum p_i (x_i - m_1) = 0,$$

et

$$\mu_2^2 = \sum p_i (x_i - m_1)^2 = m_2^2 - m_1^2.$$

On appelle μ_2 l'*écart quadratique moyen;* l'expression $x_i - m_1$ s'appelle l'*erreur*, et l'*écart* est la *valeur absolue* de l'erreur.

Une erreur est donc une variable, dont la valeur probable est nulle.

15. **Combinaison des lois de probabilité.** — Nous avons à notre disposition deux disques A et B qu'on peut faire tourner devant un index; l'un et l'autre sont divisés en secteurs de façon à représenter deux lois de probabilités différentes. J'appelle x la variable aléatoire associée au premier disque, y la variable aléatoire associée au second. A chaque valeur distincte d'une variable aléatoire correspond une probabilité déterminée.

On considère maintenant une nouvelle variable aléatoire z

$$z = f(x, y).$$

Pour obtenir une valeur de z, je détermine x en faisant tourner A, j'amène x_i qui correspond à la probabilité p_i, puis je détermine y en faisant tourner B, j'amène y_j qui correspond à la probabilité q_j. Le concours des deux événements *indépendants* qui déterminent x et y est $p_i q_j$ et la valeur z_{ij} de z est définie par

$$z_{ij} = f(x_i, y_j),$$

elle a pour probabilité $p_i q_j$.

On peut se proposer de déterminer la valeur probable de la variable aléatoire z.

Examinons, par exemple, le cas où l'on fait

$$z_{ij} = x_i y_j.$$

On aura pour la valeur probable de z que je note $\mathrm{E}(z)$

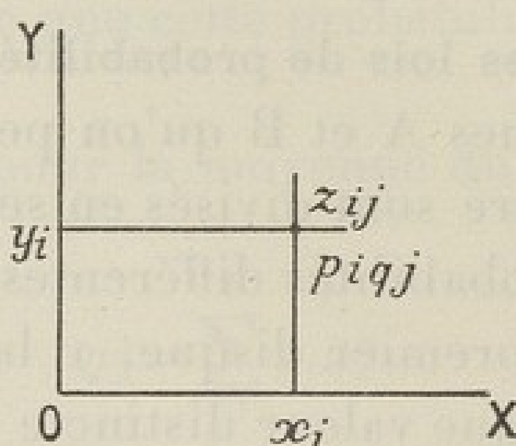
$$\mathrm{E}(z) = \sum_{i=1}^{N} \sum_{j=1}^{N'} p_i q_j x_i y_j = \sum_{i=1}^{N} p_i x_i \sum_{j=1}^{N'} q_j y_j = \mathrm{E}(x)\,\mathrm{E}(y).$$

Donc la valeur probable du produit, quand les deux variables sont *indépendantes* est égal au produit des valeurs probables.

D'une façon générale, nous pouvons représenter graphiquement la loi de la variable z, en prenant deux axes de coordonnées OX, OY; au point du plan qui a pour abscisse x_i et pour ordonnée y_j se trouve le point z correspondant z_{ij} ou (x_i, y_j) auquel est attribuée la masse $p_i q_j$. Appliquons cette représentation au cas suivant,

qui diffère du précédent en ce que les probabilités ne sont pas indépendantes.

Fig. 2.



Je détermine la valeur de x au moyen du disque A; soit x_i la valeur obtenue pour la variable aléatoire x; à x_i correspond la probabilité p_i. Dans notre nouvelle hypothèse, le disque qui sert à la détermination de la variable aléatoire y dépend du résultat obtenu pour la variable x; par exemple, admettons qu'on doive choisir dans une collection de disques B le disque B_i qui donne pour les valeurs de y à considérer, c'est-à-dire pour y_1, y_2, y_3, etc., les probabilités $q_i^{(1)}$, $q_i^{(2)}$, $q_i^{(3)}$, etc.

Alors, au point $(x_i y_j)$ correspond la masse $p_i q_i^{(j)}$, expression dans laquelle, selon la formule consacrée, $q_i^{(j)}$ est la probabilité que donne à la valeur y_j l'arrivée de l'événement $x = x_i$ pour la première variable aléatoire.

Dans ces conditions, il est bien clair que les probabilités $q_i^{(j)}$ qui correspondent à une même valeur de i, c'est-à-dire à un disque bien déterminé B_i, ont une somme égale à l'unité. Car la surface du disque est supposée égale à l'unité et les $q_i^{(j)}$ sont les valeurs numériques des surfaces des secteurs, qui représentent les probabilités correspondant aux diverses valeurs de j.

D'après le théorème de la probabilité composée, on peut écrire

$$p_i q_i^{(j)} = p_j^{(i)} q_j,$$

en appelant q_j la probabilité à priori de la valeur y_j de la variable aléatoire y et $p_j^{(i)}$ la probabilité de la valeur x_i de la variable aléatoire x, quand on est informé que y a la valeur y_j.

Il est facile de calculer q_j d'après le théorème de la probabilité totale. La valeur y_j peut être considérée comme due à diverses causes, qui sont les valeurs x_1, x_2, x_3, etc., dont les probabilités

sont p_1, p_2, p_3, etc. On a donc

$$q_j = \sum_i p_i q_i^{(j)} \quad \text{avec} \quad \sum_{(j)} q_i^{(j)} = 1.$$

On pourrait donc imaginer comme entièrement équivalent au précédent le dispositif de tirage au sort inverse. Un premier disque B′ servirait à tirer au sort la valeur de y. Ses divers secteurs auraient pour surfaces respectivement les diverses valeurs des probabilités q_j qui viennent d'être calculées. Je dis que ces valeurs sont conformes au principe de la probabilité totale et qu'on a

$$\sum_j q_j = 1,$$

on a, en effet

$$\sum_j q_j = \sum_j \sum_i p_i q_i^{(j)} = \sum_i \left(p_i \sum_j q_i^{(j)} \right) = \sum_i p_i = 1.$$

Ensuite on tirerait la valeur de x, en se servant d'un certain disque A'_j, choisi dans une collection de disques A′, ce disque correspondant à la valeur y_j obtenue au disque B′. Le disque A'_j doit être supposé divisé en secteurs représentant les probabilités des diverses valeurs de x compatibles avec l'arrivée de l'événement $(y = y_j)$. Appelons ces probabilités $p_j^{(i)}$, leur somme est égale à l'unité. L'événement $x = x_i$ pourra être considéré comme dû à diverses causes y_j dont les probabilités sont q_1, q_2, q_3, etc. On a donc pour la probabilité p_i de la proposition $x = x_i$, d'après le théorème de la probabilité totale

$$\sum_j q_j p_j^{(i)} = p_i,$$

et l'on vérifie qu'on a bien pour les probabilités p_i une somme égale à l'unité, car en vertu de ce que la somme des probabilités $p_j^{(i)}$ est égale à l'unité ainsi que $\sum_{(j)} q_i^{(j)}$, on a

$$\sum_i p_i = \sum_i \sum_j q_j p_j^{(i)} = \sum_i \left(q_j \sum_i p_j^{(i)} \right) = \sum_j q_j = 1.$$

Ces explications préliminaires étant données, cherchons la valeur probable de la variable $z = x + y$, qui résulte de la composition, par voie d'addition, des lois de probabilités, relatives à deux variables aléatoires x et y. Ici, contrairement à ce qu'on devait supposer dans le cas de $z = xy$: les deux variables, comme l'a montré M. Darmois, n'ont pas besoin d'être supposées indépendantes. On aura, en appelant p_{ij} la probabilité du couple (x_i, y_j), c'est-à-dire la masse du point (x_i, y_i)

$$\sum_i \sum_j p_{ij}(x_i + y_j) = \sum_i \sum_j p_i q_i^{(j)} x_i + \sum_i \sum_j p_j^{(i)} q_j y_j$$

$$= \sum_i \left(p_i x_i \sum_j q_i^{(j)} \right) + \sum_j \left(p_j y_j \sum_i p_j^{(i)} \right).$$

Or, on a

$$\sum_j q_i^{(j)} = 1, \qquad \sum_i p_j^{(i)} = 1.$$

D'où

$$\sum_i \sum_j p_{ij}(x_i + y_k) = \sum_i p_i x_i + \sum_j q_j y_j$$

ou, en employant une notation déjà introduite,

$$E(z) = E(x) + E(y),$$

ce qu'on peut énoncer, en disant que l'espérance mathématique relative à la variable aléatoire $z = x + y$ est égale à la somme des espérances mathématiques de x et de y.

Considérons maintenant trois lois de probabilités relatives à des variables aléatoires x, y, z dont la valeur probable moyenne est nulle. On aura donc

$$\sum_i p_i x_i = 0, \qquad \sum_i q_i y_i = 0, \qquad \sum_i r_i z_i = 0$$

et cherchons la valeur quadratique moyenne de la variable aléatoire

$$u = x + y + z,$$

p_i, q_i, r_i désignent les probabilités respectives de x_i, y_i, z_i. Nous supposons les variables *indépendantes* et ce que nous cherchons c'est la valeur du carré moyen des valeurs de u, correspondant aux

valeurs x_i, y_j, z_k où i, j, k prennent toutes les valeurs 1, 2, 3, etc. jusqu'au plus grand nombre que chacun de ces indices puisse atteindre respectivement.

On a alors à calculer

$$\sum_i \sum_j \sum_k p_i q_j r_k (x_i + y_j + z_k)^2$$

ou

$$\sum_i \sum_j \sum_k p_i q_j r_k (x_i^2 + y_j^2 + z_k^2 + 2y_j z_k + 2z_k x_i + 2x_i y_j).$$

Or, on a

$$\sum_i \sum_j \sum_k p_i q_j r_k x_i^2 = \sum_i p_i x_i^2 \sum_j q_j \sum_k r_k = \sum_i p_i x_i^2$$

et

$$\sum_i \sum_j \sum_k p_i q_j r_k y_j z_k = \sum_i p_i \sum_j q_j y_j \sum_k r_k z_k = 0.$$

D'où, en appelant p_{ijk} la masse qui est affectée au point u_{ijk} dont les coordonnées sont x_i, y_j, z_k

$$\sum p_{ijk} u_{ijk}^2 = \sum p_i x_i^2 + \sum q_j y_j^2 + \sum r_k z_k^2.$$

On peut énoncer ce résultat sous la forme suivante :

Le carré de l'écart quadratique moyen de la variable aléatoire $u = x + y + z$ est égal à la somme des carrés des écarts quadratiques moyens de x, de y et de z (ces variables étant des erreurs).

Il est clair que la propriété s'étend d'elle-même à un nombre quelconque de variables aléatoires.

Il en est encore de même et la propriété subsiste si les variables x, y, z, tout en restant indépendantes, suivent la même loi. On aura alors, par exemple,

$$\sum p_i x_i^2 = \sum q_j y_j^2 = \sum r_k z_k^2 = \mu_2^2.$$

Si le nombre des variables est égal à n et si l'on appelle σ_2 l'*écart quadratique moyen* de la variable aléatoire égale à la somme $x + y + z + \ldots$, on aura

$$\sigma_2^2 = n\mu_2^2$$

ou

$$\sigma_2 = \mu_2 \sqrt{n},$$

de sorte que si l'on considère l'écart quadratique moyen comme fonction de n, on voit qu'il croît comme la racine carrée du nombre des variables, on peut dire aussi bien des épreuves, car, puisque les lois des variables aléatoires x, y, z, etc. sont identiques l'une à l'autre, il suffit évidemment de se servir pour chacune d'elles du même disque divisé et de le faire tourner n fois, chaque fois qu'on le fait tourner constituant une épreuve ou un coup. L'écart quadratique moyen croît donc moins vite que le nombre des épreuves.

Cette loi n'est pas autre chose que la *loi des grands nombres*.

Nous avons été amenés dans l'étude des épreuves répétées, où il n'y avait en jeu que l'événement E et son contraire F avec les probabilités p et $1 - p$ respectivement, à envisager une loi de probabilité continue, que nous avons appelée la loi de Gauss. Prenons des notations très voisines de celles que nous avons adoptées précédemment et posons

$$\sqrt{2pq} = a, \qquad \lim \frac{l}{\sqrt{m}} = x,$$

nous avons pour la probabilité élémentaire

$$\frac{1}{a\sqrt{\pi}} e^{-\frac{x^2}{a^2}} dx$$

avec (*Cf.* p. 84)

$$\frac{1}{a\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{a^2}} dx = 1,$$

ainsi que nous l'avons démontré en détail. Je vais maintenant évaluer les moyennes du premier et du second ordre. En raison de la symétrie relative à l'origine, il est évident que la moyenne du premier ordre est nulle. On peut donc dire que x représente une erreur et la moyenne du second ordre sera l'écart quadratique moyen. Je l'appellerai μ_2. C'est μ_2 que je vais calculer.

La modification apportée à nos notations primitives, a pour objet de mettre en évidence la constante a, qui prend des valeurs différentes pour les différentes lois qui sont des lois de Gauss.

L'écart quadratique moyen μ_2 est relié à la constante a par

l'équation de définition

$$\mu_2^2 = \frac{1}{a\sqrt{\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{a^2}} dx.$$

Cette intégrale se calcule aisément en dérivant par rapport à a les deux membres de l'identité ci-après

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{a^2}} dx = a.$$

On obtient ainsi, par un calcul immédiat, en dérivant sous le signe somme, ce qui est permis ici, la relation

$$\mu_2^2 = \frac{a^2}{2} \quad \text{ou} \quad \mu_2 = \frac{a}{\sqrt{2}}.$$

Telle est la valeur de l'écart quadratique moyen.

16. **Théorème de d'Ocagne.** — Je passe maintenant à la démonstration du théorème connu sous le nom de théorème de d'Ocagne. Ce théorème confirme, dans le cas particulier de la loi de Gauss, la proposition précédemment établie pour l'écart quadratique moyen de la variable aléatoire $z = x + y$.

Supposons donc que x et y, de valeurs probables nulles, suivent la loi de Gauss et que la probabilité pour que x tombe dans le segment dx soit

$$\frac{1}{a\sqrt{\pi}} e^{-\frac{x^2}{a^2}} dx,$$

tandis que la probabilité pour que y tombe dans le segment dy serait

$$\frac{1}{b\sqrt{\pi}} e^{-\frac{y^2}{b^2}} dy.$$

Considérons la variable aléatoire $z = x + y$ et cherchons la probabilité que z soit compris entre z_1 et z_2. La portion du plan des x, y que nous envisageons ainsi est donc la bande illimitée comprise entre les deux parallèles

$$x + y = z_1, \qquad x + y = z_2.$$

La probabilité pour que le point (x, y) tombe dans le rectangle

élémentaire $dx\,dy$ est

$$\frac{1}{\pi ab} e^{-\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)} dx\,dy,$$

et l'on doit intégrer dans la portion du plan considérée. Il est donc naturel de rapporter l'ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

aux deux diamètres conjugués dont l'un est parallèle à la bande. Ces deux diamètres conjugués sont dirigés suivant les droites

$$x+y=0, \qquad \frac{x}{a^2}-\frac{y}{b^2}=0.$$

Mais comme elles ne sont pas rectangulaires, au lieu d'appliquer les formules du changement de coordonnées cartésiennes, nous nous bornerons à faire le changement de variables, plus simple,

$$z=x+y, \qquad u=\frac{x}{a^2}-\frac{y}{b^2}.$$

Pour faire sous le signe d'intégration le changement de variables de (x, y) en (z, u), nous avons à calculer la valeur absolue du déterminant fonctionnel, on trouve ainsi

$$\frac{\partial(z, u)}{\partial(x, y)}=-\left(\frac{1}{b^2}+\frac{1}{a^2}\right)$$

ou

$$\left|\frac{\partial(x, y)}{\partial(z, u)}\right|=\frac{a^2b^2}{a^2+b^2}.$$

Nous savons que $\frac{x^2}{a^2}+\frac{y^2}{b^2}$ doit prendre la forme

$$\frac{z^2}{A^2}+\frac{u^2}{B^2}$$

mais il est clair que A et B ne seront pas les valeurs des demi-diamètres conjugués, parce que z et u sont simplement proportionnelles aux coordonnées cartésiennes d'un point de l'ellipse.

L'identification donne

$$A^2=a^2+b^2, \qquad B^2=\frac{a^2+b^2}{a^2b^2},$$

et l'on en déduit

$$AB = \frac{a^2 + b^2}{ab}.$$

La probabilité cherchée P est donc donnée par la formule

$$P = \frac{1}{\pi} \frac{ab}{a^2 + b^2} \int_{z_1}^{z_2} \int_{-\infty}^{+\infty} e^{-\left(\frac{z^2}{A^2} + \frac{u^2}{B^2}\right)} dz\, du$$

ou, d'après le résultat précédent

$$= \frac{1}{B\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{B^2}} du \frac{1}{A\sqrt{\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{A^2}} dz.$$

Or, on a

$$\frac{1}{B\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{B^2}} du = 1,$$

d'où

$$\frac{1}{A\sqrt{\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{A^2}} dz = P \quad \text{avec } A^2 = a^2 + b^2.$$

Donc, si x et y sont deux erreurs aléatoires obéissant à la loi de Gauss, l'erreur $z = x + y$ obéit aussi à la loi de Gauss.

D'autre part, l'erreur quadratique moyenne de z a pour carré la somme de carrés des erreurs quadratiques moyennes de x et de y.

Ce dernier résultat a déjà été obtenu par une autre voie.

L'ensemble de ces propriétés s'exprime parfois en disant que la loi de Gauss est *stable*.

17. **Application à la téléphonie.** — Si un événement fortuit E est déterminé par la valeur que prend une variable x, ce même événement pourra être également bien déterminé, au moins dans un certain champ de la variable x, par la valeur que prendra une fonction $\varphi(x)$ de x. De même, si E dépend des valeurs de deux paramètres x et y, il pourra, dans un certain champ, être aussi bien déterminé par deux paramètres u et v, qui seront fonctions de x et de y. Le calcul des probabilités s'applique à des cas théoriques qui sont inventés pour être l'image aussi fidèle que possible de la réalité, mais les différents auteurs ont des idées différentes sur le mécanisme du hasard. De là viennent les controverses qui s'élèvent très souvent entre eux. Il s'agira, par exemple, de savoir

si les accroissements dx égaux ont même probabilité ou si ce sont des accroissements $d\varphi(x)$ qui jouissent de cette propriété.

Enfin, en ce qui concerne la téléphonie, on utilise les résultats de certaines statistiques; or, lorsqu'il s'agit du calcul du nombre des organes nécessaires pour desservir un trafic donné, de façon à réduire à des proportions admissibles le nombre des appels qui ne pourront être immédiatement servis, on peut se demander si les appels de cette catégorie doivent être considérés comme perdus ou comme différés et si ces pertes ou retards n'auront pas, réciproquement, une répercussion sur le trafic. Cette influence est bien difficile à évaluer, puisque ce sont des particularités qui n'étaient pas mises en évidence pendant les observations statistiques. De même, il n'est pas toujours facile de reconnaître si l'on suppose implicitement que le nombre des abonnés desservis est limité ou indéfini.

1° Supposons qu'il soit certain que, pendant l'heure la plus chargée, il se produise m appels, et considérons, au cours de cette heure, un intervalle de temps bien déterminé, dont la durée est égale à une certaine fraction p de l'heure. On demande la probabilité P, qu'il se produise α appels pendant ledit intervalle.

Nous avons déjà examiné ce problème; il nous suffit donc de rappeler qu'on a

$$P = \frac{m!}{\alpha!(m-\alpha)!}p^{\alpha}(1-p)^{m-\alpha}.$$

2° Mais il est certainement excessif de prétendre qu'il se produira juste m appels à l'heure chargée. En réalité, les statistiques ont porté sur un grand nombre d'heures, au moment de la période active du trafic journalier. On aura donc une représentation plus fidèle de la réalité en considérant, au lieu d'une heure, une période de très grande durée T. La statistique apprendra, par exemple, que dans cette durée T, il y a, en moyenne, disons un appel par unité de temps. On considère alors un intervalle de temps fini x, et l'on aurait, pour la probabilité, qu'un appel se produise pendant cet intervalle

$$p = \frac{x}{T},$$

mais comme T est infiniment grand, p est infiniment petit, et au lieu de considérer p, il vaut mieux considérer x. La durée T étant

infinie, ne peut servir d'unité. Alors, on demande la probabilité P qu'il y ait α appels dans l'intervalle de temps x (*Cf.* p. 38).

Les conditions ainsi précisées sont justement les données que nous avons admises, quand nous avons étudié la loi de Poisson. Nous nous contentons donc de rappeler la solution

$$P = \frac{x^{\alpha} e^{-x}}{\alpha!}.$$

Pour établir cette loi, nous avons admis que le nombre des conversations qui peuvent être demandées dans l'intervalle de temps x n'est pas borné supérieurement. C'est donc que le nombre des abonnés considérés est infini. D'autre part, il n'est pas question du nombre des organes ou des opératrices qui desservent ces appels. Chacun de ces appels est un appel effectif. Aucun n'est différé, ni perdu. La probabilité qu'il n'y ait pas d'appel est e^{-x}.

3° *Vie moyenne d'une conversation.* — Supposons que, le trafic battant son plein, le bureau décide de ne plus répondre aux appels et laisse les diverses conversations en cours s'acheminer vers leur fin naturelle. Soit N le nombre des conversations en cours à l'époque t après ladite décision; si nous le supposons considérable, et si le hasard seul intervient, on aura pour la variation du nombre des conversations pendant le temps dt la formule

$$dN = -aN\,dt,$$

a étant une constante. On tire de là

$$N = N_0 e^{-at}.$$

Portons N en ordonnée, t en abscisse, nous pourrons définir par la relation

$$N_0\theta = \int_0^{\infty} N_0 e^{-at}\,dt$$

une constante θ que nous appellerons la vie moyenne des conversations, et l'on a

$$\theta = \frac{1}{a}.$$

Prenons cette durée θ pour unité de temps, la formule initiale deviendra

$$dN = - N\,dt.$$

4° Dans ce paragraphe, je supposerai que θ est pris pour unité de temps.

Considérons un groupe de x lignes par lesquelles doivent s'écouler les demandes de communication, et supposons qu'il se produise en moyenne y appels par unité de temps.

On peut concevoir qu'il s'établisse, comme l'imagine M. Erlang, un *équilibre statistique* tel, qu'à l'état de régime, si P_i est la probabilité pour que, sur les x lignes, il y en ait un nombre égal à i qui soient occupées, le tableau des diverses probabilités

$$P_0,\quad P_1,\quad \ldots,\quad P_i,\quad \ldots,\quad P_x$$

demeure inchangé au cours du temps, c'est-à-dire soit indépendant de l'époque t considérée.

Nous négligeons comme infiniment peu probable le cas où plusieurs appels se produisent rigoureusement en même temps et celui où plusieurs lignes sont libérées simultanément.

Dans ces conditions, les événements qui tendent à faire passer la probabilité d'occupation des lignes de P_i à P_{i+1} doivent être compensés par ceux qui tendent à faire passer la probabilité P_{i+1} à P_i, et cela quel que soit i.

Le nombre des lignes occupées passe de i à $i+1$, s'il se produit un appel, événement dont la probabilité élémentaire est $y\,dt$, et, d'après le principe de la probabilité composée, s'il y a une probabilité P_i, pour qu'il y ait i lignes occupées, la probabilité pour que, de ce chef, on arrive à $i+1$ lignes occupées, est

$$y P_i\,dt.$$

Mais pendant ce même intervalle de temps, il y avait une probabilité P_{i+1} qu'il y eût $i+1$ lignes occupées. Ce nombre tombera à i s'il y a une conversation qui finit, événement dont la probabilité, d'après le choix qui a été fait de l'unité de temps, est $(i+1)\,dt$. Donc, la probabilité, pour que le nombre de lignes occupées tombe de $i+1$ à i lignes seulement, est

$$(i+1) P_{i+1}\,dt.$$

Comme il n'y a par hypothèse d'échange qu'entre des probabilités dont les indices diffèrent d'une unité, nous avons pour toutes les valeurs de i, de o à $x - 1$:

$$yP_i = (i+1)P_{i+1},$$

d'où

$$(1)\qquad \begin{cases} yP_0 = P_1, \\ \ldots\ldots\ldots\ldots \\ yP_{x-1} = xP_x. \end{cases}$$

D'ailleurs, comme il est certain que tous les cas possibles sont envisagés dans le tableau, on a

$$P_0 + P_1 + \ldots + P_x = 1.$$

Or, on a, évidemment,

$$y^i P_0 = i!\, P_i \quad \text{ou} \quad P_i = P_0 \frac{y^i}{i!}$$

et

$$P_0\left(1 + \frac{y}{1} + \frac{y^2}{1.2} + \ldots + \frac{y^x}{x!}\right) = 1,$$

d'où

$$P_i = \frac{\dfrac{y^i}{i!}}{1 + y + \ldots + \dfrac{y^x}{x!}},$$

ce qui est la formule d'Erlang.

Cette formule suppose que le nombre des appels est grand et que la vie moyenne des conversations est une constante, quel que soit l'état de probabilité envisagé.

On suppose qu'un observateur se rend au bureau à des intervalles de temps très courts et qu'il examine le nombre de lignes qui se trouvent occupées, il y a une probabilité P_i pour qu'il en trouve i occupées. P_x est la probabilité d'occupation de x lignes, mais la qualité du service est mieux définie par la probabilité de perte.

Voici quelques considérations.

Le nombre moyen de communications données pendant l'unité de temps est N,

$$N = P_1 + 2P_2 + 3P_3 + \ldots + xP_x.$$

Ce nombre se calcule aisément, on a d'après (1) pour y appels dans l'unité de temps

$$(2)\qquad N = y(P_0 + P_1 + \ldots + P_{x-1}) = y(1 - P_x).$$

Or, le nombre moyen de demandes de communications qui se présentent dans l'unité de temps étant égal à y, le nombre moyen de communications manquées pendant l'unité de temps est

$$y - \mathrm{N}.$$

Ce nombre est égal à $y\mathrm{P}_x$, car l'équation (2) peut se mettre sous la forme

$$(2) \qquad y - \mathrm{N} = y\,\mathrm{P}_x.$$

Donc, le rapport du nombre d'appels qu'il n'a pas été possible de servir au nombre total des appels, qui se produisent en moyenne pendant l'unité de temps, est égal à P_x.

C'est donc P_x qu'il s'agirait de rendre aussi petit que possible. Par raison d'économie, on a intérêt à diminuer x, mais l'intérêt du service est de proportionner le nombre de lignes au volume du trafic. On cherchera, par exemple, a réduire à $1/1000^{e}$ la probabilité d'un appel manqué.

Voici maintenant des considérations toutes différentes sur le même sujet.

Nous savons qu'il y a en moyenne y appels par unité de temps, et cette unité de temps a une durée égale à la durée moyenne d'une conversation. Dès lors, par application de la formule de Poisson, si nous portons notre attention sur un intervalle de temps compris entre les époques t et $t+1$, la probabilité P_i pour que i appels se produisent au cours dudit intervalle est

$$\mathrm{P}_i = e^{-y}\frac{y^i}{i!}.$$

Ouvrons une parenthèse.

Remarque. — Dans la démonstration de la formule de Poisson, j'avais supposé qu'il y avait en moyenne un point par unité de longueur et que l'intervalle considéré avait la valeur y, alors P_i est la probabilité qu'il y ait i points sur l'intervalle y. Mais on peut supposer qu'il y ait en moyenne y points par unité de longueur et que l'intervalle considéré ait la longueur 1. On trouve la même probabilité. Si N est le nombre total de points et L la longueur

totale considérée, on a dans le premier cas (*Cf.* p. 39)

$$\frac{P_{i+1}}{P_i} = \frac{y}{i+1}\frac{N-i}{L-y} \qquad \text{avec } \frac{N}{L} \rightarrow 1$$

et dans le second cas,

$$\frac{P_{i+1}}{P_i} = \frac{1}{i+1}\frac{N-i}{L-1} \qquad \text{avec } \frac{N}{L} \rightarrow y.$$

Dans l'un et l'autre cas, on a donc, quand $\frac{N}{L}$ tend vers la limite supposée

$$\lim \frac{P_{i+1}}{P_i} = \frac{y}{i+1}.$$

Il suffit de se reporter au texte de la démonstration directe, due à M. Borel, pour constater ainsi l'identité des deux résultats.

Revenons à la question.

Nous avons donc dans le tableau

$$P_0, \quad P_1, \quad P_2, \quad \ldots, \quad P_i, \quad \ldots$$

indéfiniment prolongé celui des probabilités pour que, le nombre d'abonnés et le nombre de lignes (ou organes) étant supposé illimité, il y ait, d'une façon générale, i appels qui se produisent, quel que soit t, entre les époques t et $t+1$, l'unité étant égale à la durée d'une conversation. Rien ne nous empêche de supposer les lignes de jonction numérotées et d'admettre que, pour servir un appel, on se sert toujours de la ligne disponible dont le numéro est le plus bas.

Considérons l'intervalle de temps $(t, t+1)$ et le groupe des x premières lignes.

Si, dans cet intervalle, il s'est produit i appels, les i conversations correspondantes ayant une durée égale à l'unité, occuperont i lignes à la fin de l'intervalle, et les appels qui se seront produits avant l'époque t se seront terminés avant l'époque $t+1$, donc à cette époque il y a i lignes occupées exactement.

Donc P_i est la probabilité qu'à une époque quelconque (car le régime est établi) il y ait i lignes occupées. Et, par suite, le nombre de cas où le groupe des x premières lignes aura été insuf-

fisant a pour probabilité

$$\mathrm{P} = \sum_{i=x+1}^{\infty} \mathrm{P}_i.$$

Venons-en maintenant au cas de la réalité, celui où il n'y a que x lignes. Un grand nombre de cas primitivement considérés comme possibles se trouvent écartés, et la probabilité pour qu'il y ait i lignes occupées ($i \leq x$) prend une nouvelle valeur P'_i. Alors, d'après le principe de Bayes, on a

$$\frac{\mathrm{P}_0}{\mathrm{P}'_0} = \frac{\mathrm{P}_1}{\mathrm{P}'_1} = \ldots = \frac{\mathrm{P}_i}{\mathrm{P}'_i} = \ldots = \frac{\mathrm{P}_x}{\mathrm{P}'_x},$$
$$\mathrm{P}'_0 + \mathrm{P}'_1 + \ldots + \mathrm{P}'_x = 1,$$

ce qui donne, d'après la valeur de P,

$$\frac{\mathrm{P}_0}{\mathrm{P}'_0} = \ldots = \frac{\mathrm{P}_x}{\mathrm{P}'_n} = \frac{1-\mathrm{P}}{1},$$

d'où

$$\mathrm{P}'_i = \mathrm{P}_i + \mathrm{P}\mathrm{P}'_i$$

et

$$\sum_{i=0}^{x} (\mathrm{P}'_i - \mathrm{P}_i) = \mathrm{P}.$$

On voit donc que l'on reporte la probabilité P correspondant aux appels *manqués* sur les probabilités nouvelles proportionnellement à leurs valeurs respectives.

Il ne reste plus qu'à calculer P'_i. Or, on a

$$\mathrm{P}'_i = \frac{\mathrm{P}_i}{1-\mathrm{P}} = \frac{\mathrm{P}_i}{\sum\limits_{j=0}^{x} \mathrm{P}_j} = \frac{\dfrac{e^{-y} y^i}{i!}}{\sum\limits_{j=0}^{x} \dfrac{e^{-y} y^j}{j!}} = \frac{\dfrac{y^i}{i!}}{\sum\limits_{j=0}^{x} \dfrac{y^j}{j!}},$$

ce qui est précisément la valeur

$$\mathrm{P}'_i = \frac{\dfrac{y^i}{x!}}{1 + \dfrac{y}{1} + \ldots + \dfrac{y^x}{x!}}$$

donnée par la formule d'Erlang.

D'après ces considérations, c'est la somme P des probabilités P_i

correspondant à plus de x lignes occupées qu'il conviendrait de réduire ; on se fixera, par exemple, le maximum de $1/1000^e$ pour P, et l'on augmentera x dans la mesure, qui sera juste suffisante pour obtenir ce résultat.

18. **Théorie cinétique des gaz.** — Un gaz est assimilé à un très grand nombre de sphères élastiques polies, identiques à tous égards, qui se trouvent en mouvement dans une région de l'espace de volume donné et qui, à cause de cela, entrent en collision les unes avec les autres dans toutes les directions et avec des vitesses relatives de toute grandeur. L'espace considéré est assez grand par rapport au volume total des sphères pour que le temps pendant lequel une sphère chemine sans en rencontrer une autre est infiniment plus grand que la durée d'un choc. Je dis que le nombre moyen de sphères dont les composantes de vitesse parallèles aux axes sont intermédiaires entre les valeurs

$$u \text{ et } u+du, \quad v \text{ et } v+dv, \quad w \text{ et } w+dw$$

est donné par la formule

$$\mathrm{A}\, e^{-\frac{m}{2h}(u^2+v^2+w^2)}\, du\, dv\, dw,$$

l'expression $\frac{1}{2}m(u^2+v^2+w^2)$, étant l'énergie cinétique de la sphère et m sa masse, h une constante universelle, et A ayant la valeur

En dehors des collisions, les vitesses ne changent ni de grandeur, ni de direction et la force vive ne change pas non plus; les lois de distribution ne se modifient pas. Il s'agit donc de savoir ce qui se produit au moment d'un choc.

Rappelons donc les lois du choc de deux sphères élastiques.

Prenons trois axes rectangulaires et supposons que l'axe des x soit pris dans la direction de la ligne des centres. Les sphères étant parfaitement polies, la percussion est dirigée suivant la normale. Si je désigne par les mêmes lettres, mais avec de grands caractères, les composantes des vitesses après le choc, j'ai, en appelant I la percussion subie par une sphère (m),

$$m(\mathrm{U}-u)=\mathrm{I}, \qquad m'(\mathrm{U}'-u')=-\mathrm{I},$$
$$\mathrm{V}=v, \qquad \mathrm{V}'=v',$$
$$\mathrm{W}=w, \qquad \mathrm{W}'=w'$$

et, en multipliant par $\mathrm{U}+u$ et $\mathrm{U}'+u'$ les deux premières équations, il vient

$$\frac{1}{2}m(\mathrm{U}^2-u^2)+\frac{1}{2}m'(\mathrm{U}'^2-u'^2)=\mathrm{I}\big(\mathrm{U}+u-(\mathrm{U}'+u')\big).$$

Or, la loi expérimentale du choc *élastique* dit que la composante normale de la vitesse d'approche est égale et de signe contraire à la composante normale de la vitesse de séparation après le choc. On a donc

$$\mathrm{U}-\mathrm{U}'+u-u'=0.$$

Et de là résulte la conservation de l'énergie cinétique [1]

$$\frac{1}{2}m\mathrm{U}^2+\frac{1}{2}m'\mathrm{U}'^2=\frac{1}{2}mu^2+\frac{1}{2}m'u'^2.$$

On a donc, pour trouver les valeurs des vitesses après le choc, à

(1) On peut, avec Painlevé, mener, par le centre de gravité G des deux sphères, des axes G(ξ, η, ζ) parallèles aux axes fixes. Pendant le choc, la vitesse du point G ne change pas, parce que les percussions sont intérieures; donc, d'après le théorème de Kœnig, la variation de force vive causée par le choc est égale à la variation de force vive dans le mouvement du système rapporté aux axes G(ξ, η, ζ); et l'on voit aisément que le calcul est alors réduit à sa plus simple expression.

résoudre les deux équations

$$mU + m'U' = mu + mu',$$
$$U - U' = u' - u,$$

la première exprimant la loi de conservation de la quantité de mouvement.

Pour notre but, il n'est pas nécessaire d'effectuer cette résolution; ce que nous aurons besoin de connaître c'est seulement la relation entre $dU\, dU'$ et $du\, du'$. Or d'après la règle du changement de variable ceci revient à trouver la valeur du déterminant $\frac{\partial(U, U')}{\partial(u, u')}$. Bien que ceci soit immédiat dans notre cas, cependant nous préférons recourir à un procédé plus général appliciable toutes les fois que l'énergie cinétique pourra être mise sous la forme d'une somme de carrés indépendants. Si, en effet, nous posons

$$U\sqrt{\frac{m}{2}} = X, \quad U'\sqrt{\frac{m'}{2}} = X', \quad u\sqrt{\frac{m}{2}} = x, \quad u'\sqrt{\frac{m'}{2}} = x',$$

la loi de conservation de l'énergie cinétique s'exprime simplement par la relation

$$X^2 + X'^2 = x^2 + x'^2,$$

ce qui signifie que la transformation linéaire opérée sur les variables x, x' est une transformation orthogonale. Or, on sait que le carré du déterminant d'une transformation de ce genre est égal à $+1$.

On a donc

$$dX\, dX' = dx\, dx',$$

d'où l'on déduit

$$dU\, dU' = du\, du'.$$

On peut aussi raisonner comme suit; je pose

$$mU + m'U' = \alpha,$$
$$U - U' = \beta,$$

soit Δ le déterminant fonctionnel

$$\Delta = \frac{\partial(U, U')}{\partial(\alpha, \beta)}.$$

Mais on a de même

$$mu + m'n' = \alpha,$$
$$u - u' = -\beta,$$

avec un déterminant fonctionnel Δ'

$$\Delta' = \frac{\partial(u, u')}{\partial(\alpha, \beta)} = -\Delta,$$

d'où

$$\frac{\partial(u, u')}{\partial(\mathrm{U}, \mathrm{U}')} = \frac{\Delta'}{\Delta} = -1.$$

Cette méthode est facile à généraliser.

Cela posé, le nombre des chocs entre les sphères du premier groupe dont les vitesses ont leurs composantes dans les intervalles

$$u \text{ et } u + du, \quad v \text{ et } v + dv, \quad w \text{ et } w + dw,$$

et celles du second groupe dont les vitesses ont leurs composantes dans les intervalles

$$u' \text{ et } u' + du', \quad v' \text{ et } v' + dv', \quad w' \text{ et } w' + dw',$$

pendant le temps dt pour les couples dont la ligne des centres est parallèle à l'axe des x au moment du choc, est proportionnel à

$$e^{-\frac{m}{2h}(u^2+v^2+w^2)} \, e^{-\frac{m'}{2h}(u'^2+v'^2+w'^2)} \, du \, dv \, dw \, du' \, dv' \, dw' (u - u') \, dt.$$

Le facteur $u - u'$ est la vitesse d'approche; le nombre de chocs, toutes choses égales d'ailleurs, lui est proportionnel.

Or nous venons de voir que, dans le choc élastique, la force vive est conservée. Donc, après le choc, le produit des deux exponentielles conserve la même valeur qu'avant le choc.

Les vitesses suivant l'axe des y et l'axe des z sont inchangées et l'on a aussi pour $du\,du'$ une valeur invariable; on a donc

$$du \, dv \, dw \, du' \, dv' \, dw' = d\mathrm{U} \, d\mathrm{V} \, d\mathrm{W} \, d\mathrm{U}' \, d\mathrm{V}' \, d\mathrm{W}'.$$

Enfin d'après la loi expérimentale du choc, $u - u'$ se change en $-(\mathrm{U} - \mathrm{U}')$. Ces considérations montrent que dans le choc considéré la molécule passe d'un état à un autre qu'il est possible de représenter d'une façon sensible en recourant à une image géométrique dans un hyperespace à six dimensions.

Traçons donc par la pensée les six axes rectangulaires deux à deux Ou, Ov, Ow, Ou', Ov', Ow'. Un système de deux sphères dont les composantes de vitesses tombent dans les intervalles que nous avons considérés sera représenté par un point figuratif M dans cet espace abstrait à six dimensions et ce point tombera dans le volume élémentaire $du\,dv\,dw\,du'\,dv'\,dw'$. D'après l'hypothèse faite sur la loi de distribution la densité de ces points figuratifs sera

$$e^{-\frac{T+T'}{h}},$$

T et T' étant les énergies cinétiques des deux sphères. Et nous voyons que les points figuratifs situés dans le volume élémentaire considéré vont être transportés dans un volume égal et avec la même densité. Le nombre de ces changements pendant l'unité de temps est proportionnel à $u - u'$, à l'époque considérée.

Il y a donc de ce fait une modification dans la distribution des points figuratifs entre les divers éléments de l'hyperespace considéré.

Les chocs que nous venons de spécifier sont ceux qu'on appelle chocs directs, mais étant données l'homogénéité et l'isotropie du gaz, il correspond dans la même unité de volume à chacun de ces groupes de deux sphères qui viennent d'entrer en collision et d'en sortir avec les vitesses UVW, U'V'W' un autre groupe de deux sphères qui vont entrer en collision avec des vitesses initiales égales aux vitesses finales du premier groupe. J'aurai un tel second groupe de la façon suivante : si je considère le groupe de deux sphères qui sont entrées en collision, ces deux sphéres après le choc vont s'écarter l'une de l'autre. Cependant si je remplace les deux sphères au moment du choc par les deux sphères qui leur sont symétriques par rapport au centre de gravité et si ces deux sphères dans leurs nouvelles positions sont animées des vitesses UVW, U'V'W'; alors au lieu de s'écarter elles vont entrer en collision. Or à toute combinaison de vitesses (uvw; $u'v'w'$) correspond une combinaison (UVW; U'V'W') et le nombre de collisions dans l'unité de temps est le même puisque $u - u'$ est égal à $U' - U$. Donc à chaque point figuratif qui passe du premier élément d'hypervolume dans le second correspond un point figuratif qui passe du second dans le premier. Il y a la même

densité pour les premiers que pour les seconds et le nombre de collisions est le même dans le temps dt pour les chocs du premier genre que pour les chocs du second genre. Les chocs inverses rétablissent donc la distribution qui avait été modifiée par les chocs directs.

Il résulte de cette analyse que la loi de distribution que nous avons supposée est une loi stable, en ce qui concerne les chocs des sphères (m) avec les sphères (m'). Mais rien ne nous empêche de supposer que les sphères m' soient identiques aux sphères m. Par conséquent, la distribution persiste malgré les chocs, que ceux-ci se produisent entre molécules de même espèce ou entre molécules différentes.

19. **Loi d'équipartition de l'énergie.** — Si la répartition des points figuratifs est telle que pour une certaine nature de sphères on en ait dans le volume élémentaire $du\,dv\,dw$ le nombre

$$\mathrm{A}\,e^{-\frac{m}{2h}(u^2+v^2+w^2)}\,du\,dv\,dw,$$

et s'il y en a N dans le volume réel où les sphères sont en mouvement, on doit avoir

$$\mathrm{N} = \mathrm{A}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{m}{2h}(u^2+v^2+w^2)}\,du\,dv\,dw.$$

Or, on a

$$(1) \qquad \int_{-\infty}^{+\infty} e^{-\frac{m}{2h}u^2}\,du = \sqrt{\frac{2\pi h}{m}}.$$

L'intégrale triple se décompose en un produit de trois intégrales identiques à la précédente. On a donc

$$\mathrm{N} = \mathrm{A}\left(\frac{2\pi h}{m}\right)^{\frac{3}{2}},$$

ce qui détermine A. C'est la valeur que nous avions annoncée.

Je dis maintenant que l'énergie cinétique moyenne par molécule est égale à $\frac{3}{2}h$.

Pour le démontrer, je pose

$$u^2 + v^2 + w^2 = c^2$$

et je prends comme volume élémentaire d'ingration une couche sphérique de rayon c et d'épaisseur dc; j'ai ainsi

$$du\,dv\,dw = 4\pi c^2\,dc$$

et j'ai à calculer

$$\frac{A}{N}\int_0^\infty \frac{1}{2}mc^2 e^{-\frac{m}{2h}c^2} 4\pi c^2\,dc = 2\pi m\frac{A}{N}\int_0^\infty c^4 e^{-\frac{m}{2h}c^2}\,dc;$$

or on a en dérivant deux fois la formule (1) par rapport à h (*Cf.* p. 57)

$$\int_0^\infty c^4 e^{-\frac{m}{2h}c^2}\,dc = \frac{3}{8}\sqrt{\pi}\left(\frac{2h}{m}\right)^{\frac{5}{2}},$$

il reste donc

$$2\pi m\frac{1}{(\pi)^{\frac{3}{2}}}\left(\frac{m}{2h}\right)^{\frac{3}{2}}\frac{3}{8}\sqrt{\pi}\left(\frac{2h}{m}\right)^{\frac{5}{2}} = \frac{3h}{2}.$$

Or il est évident que $\frac{1}{2}mu^2$, $\frac{1}{2}mv^2$, $\frac{1}{2}mw^2$ ont même valeur moyenne et qu'il en est de même pour le second groupe de sphères, mais qu'en outre, cette valeur moyenne ne dépend pas du groupe de sphères considéré, puisque la constante h est la même, quel que soit le groupe. On a donc pour valeur moyenne de la force vive par molécule, correspondant à chacun des trois degrés de liberté, la valeur universelle $\frac{h}{2}$.

L'extension de la démonstration à des cas de plus en plus complexes, où le nombre des degrés de liberté est de plus en plus considérable et où les chocs sont d'une nature moins simple, mais avec conservation de l'énergie, a permis d'énoncer la loi d'équipartition de l'énergie d'une façon plus générale.

La théorie des quanta a précisé son domaine d'application.

20. **Nécessité de la loi pour la répartition des vitesses.** — Considérons encore comme l'image d'un gaz monoatomique contenu dans une enceinte à une température donnée, un système mécanique formé de sphères identiques placées dans un certain volume. La distribution des sphères est uniforme dans l'espace,

mais la répartition des vitesses entre les diverses molécules s'opère suivant une certaine loi statistique. Pour avoir une représentation de ces divers états du gaz, nous représenterons chaque molécule par un point figuratif M, dans un espace idéal E qui sera le domaine de l'extension en phase. Si la molécule est animée d'une vitesse dont les composantes parallèles à trois axes de coordonnées fixes sont u, v, w, le point figuratif M aura, dans l'espace de l'extension en phase, les coordonnées u, v, w. A notre gaz réel correspondra ainsi dans cet espace idéal le brouillard des points M. La loi de répartition est définie par la formule

$$du = \mathrm{F}(u, v, w)\, du\, dv\, dw,$$

qui indique le nombre de points figuratifs M situés dans l'élément $du\, dv\, dw$. Une fois que l'équilibre statistique est établi, la fonction F est indépendante du temps. Naturellement, si N est le nombre de molécules, on doit avoir

$$\mathrm{N} = \iiint \mathrm{F}\, du\, dv\, dw. \tag{1}$$

Le nombre N de points figuratifs peut être réparti de bien des manières entre les divers éléments $d\tau = du\, dv\, dw$ de l'espace E. Il y a cependant une condition à remplir, c'est que l'énergie du gaz ne varie pas. Comme nous supposons qu'il n'y a pas d'actions des sphères les unes sur les autres en dehors des chocs, l'énergie dont il s'agit est l'énergie cinétique T. On doit donc avoir

$$\mathrm{T} = \iiint \frac{1}{2} m(u^2 + v^2 + w^2)\mathrm{F}\, d\tau = \text{const}. \tag{2}$$

Cela étant, je dis que l'on trouve pour la répartition la plus probable justement la répartition que nous avions admise précédemment et dont nous avons montré la stabilité statistique, savoir

$$\mathrm{F} = \mathrm{A}\, e^{-\frac{m}{2h}(u^2+v^2+w^2)}. \tag{3}$$

Tout d'abord, nous pouvons décomposer l'espace E en une infinité dénombrable de petits éléments de volume $d\tau_1$, $d\tau_2$, $d\tau_3$, ... Et nous supposons que chacun de ces éléments de volume, tel que $d\tau_i$ comprend un nombre très considérable dn_i de points figuratifs. Il est d'ailleurs clair que nous ne changeons pas la loi

de distribution si nous faisons permuter entre eux d'une façon quelconque ces dn_i points figuratifs. Par conséquent, nous n'avons pas à considérer toutes les N! permutations que nous pouvons faire subir entre les N points de vitesse, mais nous avons à comparer entre elles un nombre de permutations égal à W

$$W = \frac{N!}{(dn_1)!\,(dn_2)!\,(dn_3)!\,\ldots}.$$

Il est assez naturel d'admettre qu'un gaz, placé initialement dans un état quelconque, évoluera si le chaos moléculaire se maintient, vers l'état qui correspond au plus grand nombre de complexions diverses possible, c'est-à-dire vers les états les plus probables. Dans ces conditions W est pour nous une mesure de la probabilité, et c'est de cette quantité que nous allons chercher à déterminer le maximum. Pour ce calcul, comme dans l'étude des épreuves répétées, il convient de recourir à la formule de Stirling qui donne comme valeur approchée de $p\,!$ pour p très grand la valeur donnée par

$$\log(p\,!) \sim p \log p - p.$$

D'où, B étant une constante,

$$\log W = -\sum dn_i \log(dn_i) + B,$$

car

$$\sum dn_i = N$$

ou, compte tenu de (1)

$$\text{og } W = -\int F \log F \, d\tau + \text{const.}$$

Nous avons donc à chercher les conditions dans lesquelles on a

$$\delta \log W = 0 \qquad \text{ou} \qquad \int (\log F + 1)\, \delta F \, d\tau = 0$$

avec les conditions

$$\int \delta F \, d\tau = 0, \qquad \int \frac{1}{2} m(u^2 + v^2 + w^2)\, \delta F \, d\tau = 0,$$

qu'on obtient en remplaçant, dans les relations (1) et (2), F par $F + \delta F$ et faisant la différence.

On peut d'après les règles du calcul des variations, λ et μ étant deux constantes, annuler le coefficient de δF dans l'intégrale

$$\int [\log F + 1 + \lambda + \mu(u^2 + v^2 + w^2)]\, \delta F\, d\tau = 0.$$

On obtient pour F une expression de la forme (3). Les constantes A et h remplacent les constantes λ et μ, et l'on a

$$F = A\, e^{-\frac{m}{2h}(u^2+v^2+w^2)},$$

ce qui achève la démonstration. La loi de Maxwell est donc à la fois suffisante et nécessaire.

On pourrait aussi considérer une fonction analogue à W, mais relative à l'espace ordinaire et montrer que l'état le plus probable est l'état homogène.

21. Équation du mouvement brownien. — Admettons que l'énergie cinétique d'agitation moyenne E d'une molécule soit

$$E = \frac{1}{2} m\overline{c^2} = \frac{3}{2}\frac{RT}{N},$$

m étant la masse ;

c la vitesse, la barre indique qu'on doit prendre une valeur moyenne de c^2 ;

R la constante des gaz parfaits ($PV = RT$) ;

N le nombre des molécules dans une molécule-gramme ;

T la température absolue.

Observons le mouvement d'agitation de la molécule *dans la direction* de l'axe des x.

La loi du mouvement est

$$m x'' = -Kx' + X,$$

les accents indiquant des dérivations par rapport au temps ;

Kx' représente la résistance offerte par le milieu au mouvement de la molécule ;

X est la composante suivant l'axe des x des percussions dues au bombardement moléculaire.

En formant Xx, nous obtenons

$$\frac{m}{2}\frac{d^2(x^2)}{dt^2} - m\left(\frac{dx}{dt}\right)^2 = -\frac{K}{2}\frac{d(x^2)}{dt} + Xx.$$

Posons, en observant un grand nombre de particules semblables et faisant une moyenne,

$$z = \frac{d(\overline{x^2})}{dt}.$$

Observons que l'on a, pour la moyenne,

$$\overline{Xx} = 0,$$

$$\frac{1}{2} m \overline{x'^2} = \frac{RT}{2N},$$

car

$$\overline{x'^2} = \overline{y'^2} = \overline{z'^2} = \frac{1}{3}\overline{c^2},$$

l'équation peut s'écrire

$$\frac{m}{2}\frac{dz}{dt} - \frac{RT}{N} = -\frac{Kz}{2},$$

d'où, en intégrant entre 0 et τ, et C étant une constante d'intégration

$$z = \frac{2RT}{NK} + Ce^{-\frac{K}{m}\tau}.$$

Si τ est assez grand, le terme en C est négligeable, d'où

$$\frac{d(\overline{x^2})}{dt} = \frac{2RT}{NK}.$$

Si $\Delta\overline{x^2}$ est le changement de $\overline{x^2}$ dans le temps Δt, on a donc

$$\Delta\overline{x^2} = \frac{2RT}{NK}\Delta t$$

(formule d'Einstein, démonstration de Langevin).

Si donc on observe un grand nombre N de particules exactement semblables durant le temps Δt, qu'on élève au carré la valeur du déplacement subi par chacune d'elles le long de l'axe des x pendant cet intervalle de temps, et puis, qu'on prenne la moyenne, on doit obtenir la quantité

$$\frac{2RT}{NK}\Delta t.$$

Mais nous aurons le même résultat en observant la même particule pendant N intervalles de temps égaux à Δt et faisant la moyenne,

c'est préférable, car l'identité des particules est assurée, puisque c'est la même qu'on suit.

Le calcul des probabilités se réduit le plus souvent en physique mathématique à un calcul de moyennes, comme on le voit par cet exemple.

Démonstration sommaire de la formule (1)

$$(1) \qquad \frac{1}{2} m\overline{c^2} = \frac{3}{2} \frac{RT}{N}.$$

Comme il s'agit, en définitive, d'une comparaison avec la loi des gaz parfaits

$$PV = RT \qquad (\text{V, volume de la molécule-gramme})$$

c'est la pression du gaz qu'il s'agit de calculer.

La pression P du gaz sur une paroi de l'enceinte est due au choc élastique des molécules sur l'unité de surface S.

Considérons une molécule de masse m, animée d'une certaine vitesse $\vec{c}(u, v, w)$, déterminée en sens, grandeur et direction; elle rebondit sur la surface S en lui cédant une quantité de mouvement $2mu$, si u est la composante normale. Considérons de même toutes les molécules de vitesse équipollente à $\vec{c}$ et qui dans le temps dt viennent frapper S; elles sont comprises dans un cylindre oblique parallèle à $\vec{c}$, de base S et de hauteur $u\,dt$. Si le nombre de ces molécules par unité de volume est ν, il y en aura $\nu u\,dt$ et la quantité de mouvement perdue par le gaz dans l'unité de temps au voisinage de S sera $2\Sigma\nu mu^2$, en étendant la sommation à tous les groupes de molécules qui, animées de vitesses diverses $\vec{c}$ viennent frapper la paroi, et qui sont au nombre de ν dans chaque groupe, ce nombre ν pouvant varier suivant la vitesse considérée. Introduisons le carré moyen en observant que sur les n molécules que contient le gaz par unité de volume, la moitié seulement s'approche de la paroi, les autres s'en écartant. Il faut donc envisager $\frac{n}{2}$ molécules seulement et tenir compte de la relation

$$(\overline{u^2}) = \frac{1}{3}(\overline{c^2}).$$

Il vient ainsi, puisque la pression est égale à la résultante des impulsions considérées,

$$P = \frac{1}{3} n m \overline{c^2},$$

d'où, V étant le volume de la molécule-gramme

$$\frac{1}{3} m \overline{c^2} = \frac{RT}{Vn} = \frac{RT}{N}.$$

C'est la formule, que nous nous proposions d'établir.

Bibliographie. — Sur le calcul du nombre des organes entrant dans la constitution d'un bureau téléphonique automatique, consulter le *Cours de téléphonie automatique*, livre II, de M. Ch. Petit, Ingénieur en Chef des Postes et Télégraphes (1932) et les Mémoires originaux de Molina, Merker, Vaulot, Engset, Pollaczeck, etc.

NOTE.

VALEUR ASYMPTOTIQUE DE n !

D'après la définition de la fonction gamma, on a

$$\Gamma(n+1) = \int_0^\infty x^n e^{-x}\, dx.$$

Je dis que quand n est entier, on a :

$$\Gamma(n+1) = n!$$

Pour le voir, il suffit de former la série en t

$$S = \sum_0^\infty \Gamma(n+1)\frac{t^n}{n!}.$$

On a, en effet, d'après la représentation de $\Gamma(n+1)$ par une intégrale

$$S = \int_0^\infty \sum_0^\infty \frac{x^n t^n}{n!} e^{-x}\, dx = \int_0^\infty e^{-x(1-t)}\, dx = \frac{1}{1-t}.$$

D'où

$$S = \sum_0^\infty t^n.$$

En comparant les coefficients de t^n dans les deux expressions de S, on a la relation annoncée :

$$\frac{\Gamma(n+1)}{n!} = 1 \qquad (n, \text{ entier}).$$

Considérons donc, pour avoir une valeur approchée de n ! la fonction $\Gamma(n+1)$.

Si je pose

$$y = x^n e^{-x},$$

je dis que y est maximum pour $x = n$ dans l'intervalle de o à l'infini; c'est ce qui est évident quand on prend la dérivée logarithmique

$$\frac{y'}{y} = \frac{n}{x} - 1 \qquad (x > 0;\ y > 0),$$

car quand x varie de zéro à $+\infty$, y' passe du positif au négatif pour $x = n$, valeur pour laquelle y' s'annule et qui rend y maximum.

La courbe qui représente y en fonction de x part de l'origine, s'élève au-dessus de l'axe des x, atteint un maximum pour $x = n$ et décroît ensuite jusqu'à zéro pour x infini.

C'est également l'allure de la courbe

$$y = n^n e^{-n} e^{-v^2},$$

quand v croît de $-\infty$ jusqu'à $+\infty$. Elle part tangentiellement à l'axe des v et y revient, atteignant son maximum qui a lieu pour $v = 0$ et qui a la même valeur que celui de la courbe précédente.

Nous établirons une correspondance entre les points de même ordonnée des deux courbes; quand x va de zéro à n, la variable v ira de $-\infty$ à zéro et quand x va de n à $+\infty$, la variable v ira de zéro à $+\infty$.

Et nous poserons

$$n^n e^{-n} e^{-v^2} = x^n e^{-x}.$$

Effectuons le changement de variable dans l'intégrale. Elle devient

$$\int_{-\infty}^{+\infty} n^n e^{-n} e^{-v^2} \frac{dx}{dv}\, dv,$$

et nous avons à calculer $\frac{dx}{dv}$. Or on a

$$n \log x - x = n \log n - n - v^2,$$

d'où, en différentiant,

$$\left(\frac{n}{x} - 1\right) dx = -2v\, dv$$

et, en posant

$$x = n + z,$$

on a

$$\frac{dx}{dv} = 2v\left(1 + \frac{n}{z}\right).$$

Par la même substitution, la relation entre x et v prend la forme

$$n \log(n + z) - n - z = n \log n - n - v^2.$$

On ne peut songer à développer en *série* suivant les puissances entières

de z la fonction $\log\left(1+\frac{z}{n}\right)$, puisque z est susceptible de prendre des valeurs plus grandes que n, mais on peut appliquer à $\log(n+z)$ la formule de Taylor limitée au second terme. On a ainsi

$$\log(n+z) = \log n + \frac{z}{n} - \frac{1}{2} z^2 \frac{1}{(n+\theta z)^2},$$

θ étant une fonction de z dont on sait que la valeur est comprise entre zéro et 1.

Reportons $\log(n+z)$ dans l'expression précédente, elle donne

$$v^2 = \frac{nz^2}{2} \frac{1}{(n+\theta z)^2}.$$

Nous pouvons extraire la racine carrée et choisir le signe en nous rappelant que v et z sont toujours de même signe. On a ainsi

$$2v = \sqrt{2n}\,\frac{z}{n+\theta z}.$$

Or z figure dans l'expression de $\frac{dx}{dv}$ par l'expression $1+\frac{n}{z}$. Le calcul au moyen de l'expression précédente donne

$$1+\frac{n}{z} = 1 - \theta + \frac{\sqrt{2n}}{2v}.$$

Je supposerai maintenant que, dans la fonction θ de z, j'ai remplacé z par sa valeur en fonction de v. Cette substitution n'empêche pas que θ et par suite $1-\theta$ reste toujours compris entre zéro et 1.

J'ai donc

$$\frac{dx}{dv} = \sqrt{2n} + (1-\theta)2v$$

et l'intégrale devient

$$\Gamma(n+1) = n^n e^{-n} \sqrt{2n} \int_{-\infty}^{+\infty} e^{-v^2} \left(1 + \frac{(1-\theta)}{\sqrt{2n}} 2v\right) dv.$$

On a

$$\int_{-\infty}^{+\infty} e^{-v^2}\, dv = \sqrt{\pi}.$$

Et il est facile d'avoir une limite supérieure de la valeur de l'intégrale restante

$$I = \frac{1}{\sqrt{2n}} \int_{-\infty}^{+\infty} e^{-v^2} (1-\theta)\, 2v\, dv,$$

θ n'a pas la même valeur pour deux valeurs de v égales et de signes

contraires, mais dans tous les cas $1-\theta$ est compris dans l'intervalle de zéro à 1. L'intégrale I est la somme de deux autres, si l'on décompose le champ d'intégration en deux intervalles allant de $-\infty$ à zéro et de zéro à $+\infty$. Tous les éléments de la première sont négatifs, tous les éléments de la seconde sont positifs.

La seconde est inférieure à $\frac{1}{\sqrt{2n}}$, car on a

$$\frac{1}{\sqrt{2n}}\int_0^\infty e^{-v^2}(1-\theta)\,2v\,dv < \frac{1}{\sqrt{2n}}\int_0^\infty e^{-v^2}\,2v\,dv = \frac{1}{\sqrt{2n}}.$$

En changeant v en $-v$ dans la première θ se changera en une autre fonction θ', mais on a toujours $0<\theta'<1$ et l'on verrait que la valeur absolue de l'intégrale est aussi inférieure à $\frac{1}{\sqrt{2n}}$.

L'intégrale I qui est la différence de deux nombres inférieurs à $\frac{1}{\sqrt{2n}}$ est elle-même inférieure en valeur absolue à $\frac{1}{\sqrt{2n}}$. On peut donc écrire, ω étant inférieur à 1 en valeur absolue :

$$\Gamma(n+1) = n^n e^{-n}\sqrt{2\pi n}\left(1+\frac{\omega}{\sqrt{2\pi n}}\right) \qquad (-1<\omega<+1).$$

Telle est la formule de Stirling. Le rapport de $\Gamma(n+1)$ à la valeur asymptotique $n^x e^{-n}\sqrt{2\pi n}$ tend vers l'unité quand n augmente indéfiniment. Cette méthode est due à M. Jordan.

L'emploi de la valeur asymptotique à la place de $n!$ entraîne une erreur absolue qui peut être grande, mais l'erreur relative tend vers zéro, quand n augmente indéfiniment.

10. Valeur de l'intégrale

$$I = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx.$$

Je dis que l'on a

$$I = \lim \int_{-a}^{+a} e^{-\frac{x^2}{2}}\,dx \qquad \text{pour } a \to \infty.$$

En effet, quand le champ $2a$ augmente, la valeur de l'intégrale va en croissant. Or on a

$$\int_{-a}^{+a} e^{-\frac{x^2}{2}}\,dx < 2\sqrt{2}\int_0^\infty e^{-u}\,du = 2\sqrt{2}\left|\frac{e^{-u}}{-1}\right|_0^\infty = 2\sqrt{2}.$$

Ainsi cette valeur reste toujours inférieure à une borne fixe $2\sqrt{2}$. Elle

converge donc vers une valeur finie, ce qui démontre que l'on a bien pour I une valeur positive finie et déterminée.

On peut écrire de même en appelant y la variable d'intégration

$$\mathrm{I} = \lim \int_{-a}^{+a} e^{-\frac{y^2}{2}}\, dy$$

et par suite

$$\mathrm{I}^2 = \lim \int_{-a}^{+a} \int_{-a}^{+a} e^{-\frac{x^2+y^2}{2}}\, dx\, dy.$$

Le champ d'intégration est le carré dont le centre est à l'origine et qui est compris entre les droites

$$x = \pm a, \qquad y = \pm a.$$

Le champ d'intégration est plus grand que la surface intérieure S_1 du cercle de rayon a et plus petit que la surface intérieure S_2 du cercle de rayon $a\sqrt{2}$, le premier étant le cercle inscrit dans le carré, le second le cercle circonscrit à ce même carré, et il est évident que l'on a

$$\text{im} \int_{S_1} e^{-\frac{x^2+y^2}{2}}\, dx\, dy < \mathrm{I}^2 < \lim \int_{S_2} e^{-\frac{x^2+y^2}{2}}\, dx\, dy.$$

Je décompose alors l'aire S_1 ou l'aire S_2 en couronnes concentriques de rayon r, la surface élémentaire étant égale à $2\pi r\, dr$. On a alors

$$\lim \int_0^a e^{-\frac{r^2}{2}}\, 2\pi r\, dr < \mathrm{I}^2 < \lim \int_0^{a\sqrt{2}} e^{-r^2}\, 2\pi r\, dr$$

ou

$$\lim 2\pi\left(1 - e^{-\frac{a^2}{2}}\right) < \mathrm{I}^2 < \lim 2\pi(1 - e^{-a^2}) \qquad \text{pour } a \to \infty,$$

d'où

$$\mathrm{I} = \sqrt{2\pi}.$$

Telle est la valeur de l'intégrale cherchée.



TABLE DES MATIÈRES.

PRINCIPES GÉNÉRAUX.

Pages.
1. Notion de probabilité 1
2. Probabilité composée 6
3. Événement indépendant 9
4. Probabilité totale 11
5. Principe de Bayes 15
6. Application du Calcul des probabilités à l'analyse combinatoire 19

SUR LA RÉPÉTITION DES ÉPREUVES ET LE PRINCIPE DE BAYES.

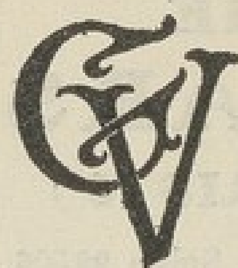
7. Épreuves répétées 25
8. Lois de probabilités continues (Loi de Gauss) 35
9. Loi de Poisson 37
Représentation géométrique de la loi de poisson 38
10. Probabilités géométriques 40
11. Problème de la Poule 43
12. Origine du Calcul des probabilités 45
13. Utilité des méthodes de récurrence 47
14. Variable aléatoire. Coefficients caractéristiques. Loi des grands nombres. 47
15. Combinaison des lois de probabilité 51
16. Théorème de d'Ocagne 57
17. Téléphonie. Calcul du nombre des organes 59
18. Théorie cinétique des gaz. Équipartition de l'énergie 67
19. Loi d'équipartition de l'énergie 72
20. Nécessité de la loi pour la répartition des vitesses 73
21. Mouvement brownien 76

BIBLIOGRAPHIE 79

NOTE : Valeur asymptotique de n! 81

TABLE DES MATIÈRES 87

BIBLIOTHÈQUE NATIONALE R.F. IMPRIMÉS



LIBRAIRIE-IMPRIMERIE
GAUTHIER-VILLARS
55, Quai des Grands-Augustins, PARIS (6e)
Tél. DANTON 05-11 et 05-12. R. C. Seine 99506.

Envoi dans toute la France et l'Union postale contre mandat-poste ou valeur sur Paris.
(Chèques postaux : Paris 29323.)
Frais de port en sus : France et France d'Outre-Mer, 5 %; Étranger, 10 %.

Paul LÉVY

Professeur à l'École Polytechnique

Calcul des Probabilités



In-8 (25-16) de VIII-350 pages, avec figures; 1925 . . . . 50 fr.



LIBRAIRIE-IMPRIMERIE
GAUTHIER-VILLARS
55, Quai des Grands-Augustins, PARIS (6e)
Tél. DANTON 05-11 et 05-12. R. C. Seine 99506.

Envoi dans toute la France et l'Union postale contre mandat-poste ou valeur sur Paris.
(Chèques postaux : Paris 29323.)
Frais de port en sus : France et France d'Outre-Mer, 5 °/o ; Étranger, 10 °/o.

R. DE MONTESSUS DE BALLORE
Docteur ès sciences mathématiques
Lauréat de l'Institut

Leçons élémentaires

sur le

Calcul des Probabilités

Philosophie du hasard. Principe du calcul des probabilités. Jeux du hasard. Jeux savants. La spéculation. Probabilité géométrique. Probabilité des causes. Tir des armes à feu. Les assurances. Les sciences morales et économiques.

In-8 (25-16) de VI-190 pages, avec 17 figures ; 1908 . . . . 20 fr.



LIBRAIRIE-IMPRIMERIE
GAUTHIER-VILLARS
55, Quai des Grands-Augustins, PARIS (6e)
Tél. DANTON 05-11 et 05-12. R. C. Seine 99506.

Envoi dans toute la France et l'Union postale contre mandat-poste ou valeur sur Paris.
(Chèques postaux : Paris 29 323.)
Frais de port en sus : France et France d'Outre-Mer, 5 %; Étranger, 10 %.

H. POINCARÉ
Membre de l'Institut.

Calcul des Probabilités

Rédaction de A. QUIQUET
Ancien Élève de l'École Normale supérieure

DEUXIÈME ÉDITION REVUE ET AUGMENTÉE

In-8 (25-16) de 336 pages; 1912. Nouveau tirage; 1923. . . . . 32 fr.

102660 Paris. — Imprimerie GAUTHIER-VILLARS, quai des Grands-Augustins, 55

TRAITÉ DU CALCUL DES PROBABILITÉS ET DE SES APPLICATIONS

Collection d'Ouvrages sur le Calcul des Probabilités

Publiée par **Émile BOREL**, Professeur de Calcul des Probabilités et Physique mathématique à la Faculté des Sciences de Paris, Membre de l'Institut.

Avec la collaboration de C.-V.-L. CHARLIER, R. DELTHEIL, P. DUBREIL, M. FRÉCHET, H. GALBRUN, J. HAAG, R. LAGRANGE, F. PERRIN, RISSER, C.-E. TRAYNARD.

Volumes in-8 raisin (25-16) se vendant séparément :

TOME I. — **Les principes de la Théorie des probabilités.**

1. Principes et formules classiques, par E. BOREL, rédigé par R. LAGRANGE. **25 fr.**
2. Erreurs et moindres carrés, par Robert DELTHEIL .... **30 fr.**
3. Recherches théoriques modernes, par M. FRÉCHET.
4. Les principes de la statistique mathématique, par R. RISSER et C.-E. TRAYNARD .... **80 fr.**

TOME II. — **Les applications de la Théorie des probabilités aux sciences mathématiques et aux sciences physiques.**

1. Applications à l'arithmétique et à la théorie des fonctions, par Émile BOREL, rédigé par DUBREIL .... **25 fr.**
2. Probabilités géométriques, par Robert DELTHEIL .... **30 fr.**
3. Mécanique statistique classique, par Émile BOREL, rédigé par Francis PERRIN .... **25 fr.**
4. Applications à l'astronomie, par C.-V.-L. CHARLIER .... **40 fr.**
5. Applications aux théories physiques actuelles, par Émile BOREL et Francis PERRIN.

TOME III. — **Les applications de la Théorie des probabilités aux sciences économiques et aux sciences biologiques.**

1. Assurances sur la vie. Calcul des primes, par Henri GALBRUN .... **45 fr.**
2. Assurances sur la vie. Calcul des réserves, par Henri GALBRUN ..... **40 fr.**
3. Applications de la statistique à la démographie et à la biologie, par R. RISSER .... **50 fr.**
4. Théorie mathématique de l'assurance invalidité et de l'assurance nuptialité. Définitions et relations fondamentales, par Henri GALBRUN .... **40 fr.**
5. Théorie mathématique de l'assurance invalidité et de l'assurance nuptialité. Calcul des primes et des réserves, par Henri GALBRUN .... **45 fr.**
6. Théorie mathématique de l'assurance maladie, par Henri GALBRUN. **60 fr.**

TOME IV. — **Applications diverses et conclusion.**

1. Applications au tir, par J. HAAG .... **35 fr.**
2. Applications aux jeux de hasard, par Émile BOREL.
3. Compléments divers.
4. Conclusion : la portée philosophique de la Théorie des probabilités, par Émile BOREL.

102660-35. Paris. — Imp. Gauthier-Villars, 55, quai des Grands-Augustins (6e).


BIBLIOTHEQUE NATIONALE DE FRANCE
3 7531 05741510 2